Robust radiotherapy planning

Unkelbach, Jan ; Alber, Markus ; Bangert, Mark ; Bokrantz, Rasmus ; Chan, Timothy C Y ; Deasy, Joseph O ; Fredriksson, Albin ; Gorissen, Bram L ; van Herk, Marcel ; Liu, Wei ; Mahmoudzadeh, Houra ; Nohadani, Omid ; Siebers, Jeffrey V ; Witte, Marnix ; Xu, Huijun

Abstract: Motion and uncertainty in radiotherapy is traditionally handled via margins. The clinical target volume (CTV) is expanded to a larger planning target volume (PTV), which is irradiated to the prescribed dose. However, the PTV concept has several limitations, especially in proton therapy. Therefore, robust and probabilistic optimization methods have been developed that directly incorporate motion and uncertainty into treatment plan optimization for intensity modulated radiotherapy (IMRT) and intensity modulated proton therapy (IMPT). Thereby, the explicit definition of a PTV becomes obsolete and treatment plan optimization is directly based on the CTV. Initial work focused on random and systematic setup errors in IMRT. Later, inter-fraction prostate motion and intra-fraction lung motion became a research focus. Over the past ten years, IMPT has emerged as a new application for robust planning methods. In proton therapy, range or setup errors may lead to dose degradation and misalignment of dose contributions from different beams - a problem that cannot generally be addressed by margins. Therefore, IMPT has led to the first implementations of robust planning methods in commercial planning systems, making these methods available for clinical use. This paper first summarizes the limitations of the PTV concept. Subsequently, robust optimization methods are introduced and their applications in IMRT and IMPT planning are reviewed.

DOI: https://doi.org/10.1088/1361-6560/aae659

Robust radiotherapy planning

Jan Unkelbach[1], Markus Alber[2,3], Mark Bangert[3,4], Rasmus Bokrantz[5], Timothy CY Chan[6], Joseph O Deasy[7], Albin Fredriksson[5], Bram L Gorissen[8], Marcel van Herk[9], Wei Liu[10], Houra Mahmoudzadeh[11], Omid Nohadani[12], Jeffrey V Siebers[13], Marnix Witte[14], Huijun Xu[15]

[1] Department of Radiation Oncology, University Hospital Zürich, Switzerland
[2] Department of Radiation Oncology, University Hospital Heidelberg, Germany
[3] Heidelberg Institute for Radiation Oncology (HIRO), Heidelberg, Germany
[4] Department of Medical Physics in Radiation Oncology, German Cancer Research Center (DKFZ), Heidelberg, Germany
[5] RaySearch Laboratories, Stockholm, Sweden
[6] Department of Mechanical and Industrial Engineering, University of Toronto, Canada
[7] Department of Medical Physics, Memorial Sloan-Kettering Cancer Center, New York, NY, USA
[8] Department of Radiation Oncology, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA
[9] Institute of Cancer Sciences, University of Manchester, Manchester, UK
[10] Department of Radiation Oncology, Mayo Clinic Arizona, Phoenix, AZ, USA
[11] Department of Management Sciences, University of Waterloo, Canada
[12] Department of Industrial Engineering & Managment Sciences and Department of Radiation Oncology, Northwestern University, Evanston, IL, USA
[13] Department of Radiation Oncology, University of Virginia, Charlottesville, VA, USA
[14] Department of Radiation Oncology, The Netherlands Cancer Institute, Amsterdam, NL
[15] Department of Radiation Oncology, University of Maryland School of Medicine, Baltimore, MD, USA

E-mail: `jan.unkelbach@usz.ch`

Abstract. Motion and uncertainty in radiotherapy is traditionally handled via margins. The clinical target volume (CTV) is expanded to a larger planning target volume (PTV), which is irradiated to the prescribed dose. However, the PTV concept has several limitations, especially in proton therapy. Therefore, robust and probabilistic optimization methods have been developed that directly incorporate motion and uncertainty into treatment plan optimization for intensity modulated radiotherapy (IMRT) and intensity modulated proton therapy (IMPT). Thereby, the explicit definition of a PTV becomes obsolete and treatment plan optimization is directly based on the CTV. Initial work focused on random and systematic setup errors in IMRT. Later, inter-fraction prostate motion and intra-fraction lung motion became a research focus. Over the past 10 years, IMPT has emerged as a new application for robust planning methods. In proton therapy, range or setup errors may lead to dose degradation and misalignment of dose contributions from different beams a problem

that cannot generally be addressed by margins. Therefore, IMPT has led to the first implementations of robust planning methods in commercial planning systems, making these methods available for clinical use. This paper first summarizes the limitations of the PTV concept. Subsequently, robust optimization methods are introduced and their applications in IMRT and IMPT planning are reviewed.

1. Introduction

Radiotherapy aims at delivering curative doses of radiation to tumors while minimizing the risk of side effects in healthy tissues. In that regard, radiotherapy treatment planning and delivery faces many uncertainties. Target volume definition, the first step in the treatment planning chain, is associated with substantial uncertainty. Definition of the gross tumor volume (GTV) has limitation not only due to finite resolution of medical images, but also because current imaging modalities only visualize surrogates for the presence of tumor and not tumor cells per se. Delineation of the clinical target volume (CTV) faces even larger uncertainty because currently available imaging modalities cannot visualize microscopic disease. Subsequently, there is uncertainty in dose prescription and normal tissue tolerances. For an individual patient, the dose that is needed to control the tumor is uncertain. Current research aims to predict an individual patient's response to radiation based on biomarkers in order to personalize prescription doses or normal tissue constraints, however, such approaches are not yet widely established. In summary, the ideal dose distribution that radiotherapy planning should be aiming at is uncertain in the first place.

In addition, there is uncertainty in the dose distribution delivered to the patient, i.e. potential discrepancies between the dose distribution shown in the treatment planning system and the actually delivered dose. The most prominent reasons for that are setup errors, changes of the patient geometry over the course of treatment, and uncertainty in dose calculation. Changes in the patient geometry include, for example, inter-fraction motion of the prostate as well as intra-fraction motion in the lung or liver due to respiration. Dose calculation errors arise in part from the use of approximate pencil beam dose calculation algorithms, which are used for computational efficiency at the cost of lower accuracy compared to methods that are directly based on modeling the physical interactions of radiation in tissue. In proton therapy, range uncertainty can be considered a specific form of dose calculation uncertainty. The Hounsfield numbers of the planning CT are in unideal input for proton dose calculation algorithms due to uncertainty in the conversion of Hounsfield numbers to proton stopping power. In addition, pencil beam dose calculation algorithms may inaccurately model the degradation of the Bragg peak in heterogeneous media.

This paper will deal with uncertainties in the delivered dose distribution. Treatment planning should aim at creating plans that are robust against uncertainty. Robustness of

a treatment plan refers to two properties: first, the CTV should receive the prescribed dose despite errors that may occur; and second, normal tissue constraints should be satisfied despite potential errors in planning or delivery. Setup and motion-related uncertainty is traditionally handled via safety margins, i.e. by expanding the irradiated region around the tumor. In IMRT planning, margins are added around the CTV in order to obtain the planning target volume (PTV). Treatment planning aims to have the PTV receive the prescription dose. It is then assumed that, as long as the CTV moves only within the boundaries of the PTV, the prescribed dose is delivered to the CTV. The required margin depends on the magnitude of the error and general margin recipes have been developed [1, 2, 3]. Typically, the priority in treatment planning is to make sure that the CTV receives the prescribed dose despite uncertainty. In specific cases, especially when the OAR is serial, respecting maximum dose constraints to normal tissues is prioritized over CTV coverage. An example is stereotactic body radiotherapy for spinal metastases, where sparing of the spinal cord is more important than target coverage. In this case, the spinal cord is expanded by a margin to create a planning risk volume (PRV). Treatment planning creates a plan that does not exceed the maximum dose to the spinal cord in all of the PRV.

The PTV concept has several limitations. To address these limitations, robust planning methods have been developed that directly incorporate uncertainty into treatment plan optimization for IMRT and IMPT. Thereby, the definition of a PTV or PRV becomes obsolete and treatment planning is based on the CTV directly. In September 2015, the authors and other researchers met at Massachusetts General Hospital in Boston to discuss the state-of-the-art in robust treatment planning. The goals of this joint review are:

1. We first summarize the limitations of the PTV concept to provide the motivation for robust planning (section 2).
2. We formally introduce the main concepts used in robust planning, namely stochastic programming and minimax optimization (section 3).
3. We review the main applications of robust planning. In particular, random and systematic setup errors and inter-fraction organ motion in IMRT (section 4), systematic range and setup errors in IMPT (section 5), and respiratory motion (section 6).

We provide a comprehensive review of robust planning approaches found in the literature. Different approaches are formulated using a unified notation. Thereby, their relation and differences can be understood. This paper serves as both a review of the literature as well as a tutorial style introduction to the concepts and main applications of robust optimization in radiotherapy.

2. Limitations of the PTV concept

While margins and PTVs are used throughout in clinical practice, there are several limitations of the PTV concept, which motivate the development of robust planning. In some situations, the PTV concept has fundamental limitations and does not guarantee target coverage irrespective of the size of the margin; in other situations a large enough PTV may ensure coverage of the CTV, but may not yield the optimal tradeoff between target coverage and OAR sparing.

1. **Breakdown of the static dose cloud approximation:** The PTV concept as typically applied in IMRT planning relies on the so-called static dose cloud approximation, i.e. the assumption that the dose distribution in treatment room coordinates is unaffected by changes in the patient's anatomy. That is, it is assumed that the CTV receives the prescribed dose as long as it stays within the PTV. This fundamental assumption is not generally fulfilled and is violated, in particular, in IMPT (section 5).
2. **Build-in margins for non-conformal plans:** Whether or not the CTV receives the prescribed dose depends on the dose distribution rather than geometric margin concepts. In reality, dose distributions are neither perfectly conformal to the PTV nor equally conformal on all sides of the CTV. Non-conformity results in an inherent dosimetric margin [4]. In those regions where the prescription isodose line extends beyond the CTV anyway, less or no margin needs to be added to account for setup uncertainty. In addition to conformity, the required margin also depends on the steepness of the dose fall-off near the target. A naturally shallow fall-off may require a smaller margin than a steep fall-off. The optimal margin may therefore be unisotropic.
3. **Optimally balancing tumor coverage and normal tissue sparing:** TCP and NTCP models are increasingly used for treatment plan evaluation, and may play a larger role in treatment plan optimization in the future. However, using a PTV dose distribution as input to a TCP model has conceptual flaws. Furthermore, underdosing the edge of the PTV may give low predicted TCP values, even though the underdosage may only occur for specific setup errors in the corresponding direction, while in most cases the CTV is covered. Optimally balancing TCP and NTCP requires proper handling of geometric uncertainties.
4. **Dose painting:** In the context of dose painting based on functional imaging, the use of margins for different prescription dose levels becomes cumbersome.
5. **Edge enhancement or horns:** The PTV approach, as well as the internal target volume (ITV) approach for respiratory motion aim to deliver the prescribed dose to all regions were the tumor may be. Thereby, the same dose is delivered to regions that are always occupied by tumor and regions where the tumor is rarely. In the presence of motion, and when the total tumor dose is achieved by accumulating dose contributions from multiple geometric instances, the approach is suboptimal in

terms of normal tissue sparing. The normal tissue dose can be reduced by delivering less dose to regions where the tumor is rarely. In order to not compromise target coverage, this has to be compensated for by delivering higher doses to regions mostly occupied by the tumor. Such dose distributions, which show dose hot spots inside or at the edge of the target when delivered to a static geometry, have been referred to as *edge enhanced* or *horns* (sections 4.2, 6).

3. Robust planning concepts

3.1. Conventional treatment plan optimization

IMRT and IMPT treatment planning is generally formulated as a mathematical optimization problem. Treatment plan quality is mathematically defined via an objective function f. A *good* treatment plan corresponds to a low objective function value. The best treatment plan is found by minimizing the objective function with respect to beamlet intensities using mathematical optimization algorithms. The objective function $f(d;q)$ is a function of the dose distribution d, and additionally depends on parameters q such as prescription doses and tolerances for normal tissues. Formally, the fluence map optimization problem for IMRT and IMPT can be written as

$$\underset{x}{\text{minimize}} \quad f(d;q) \tag{1}$$

$$\text{subject to} \quad d = Dx \tag{2}$$

$$x \geq 0 \tag{3}$$

The dose distribution $d = Dx$ is a linear function of the incident fluence x. D denotes the dose-influence matrix whose elements D_{ij} store the dose contribution of beamlet j to voxel i for unit fluence.

3.2. Types of uncertainty

In this setting, three types of uncertainty can be distinguished.

1. Uncertainty in the dose-influence matrix D. This means that the same treatment plan, as defined via its incident fluence x, may lead to different dose distributions in the patient. Uncertainty in D models geometric uncertainty such as setup errors, organ motion, and range errors.
2. Uncertainty in the realized fluence x. This would mean that the treatment machine does not accurately deliver the fluence specified by the treatment plan. This uncertainty is typically considered small compared to uncertainty in the dose-influence matrix since the accuracy of fluence delivery can be verified during machine QA. The work by Bertsimas et al [5] applies robust optimization to uncertainty in the realized fluence.
3. Uncertainty in q. This can be interpreted as uncertainty in the mathematical definition of what a good dose distribution is. q could, for example, represent

uncertain parameters in a TCP or NTCP model, uncertainty in target delineation, or uncertainty in dose prescription. This is generally considered a large source of uncertainty, however, convincing applications of robust optimization in that context are limited.

Robust optimization applied to geometric uncertainty, i.e. uncertainty in the dose-influence matrix D has been studied most widely and is the focus of this review. For simplicity of notation we assume that geometric uncertainty is modeled through discrete error scenarios indexed by k. Each error scenario corresponds to a different dose-influence matrix D^k, which yield distinct dose distributions

$$d^k = D^k x \tag{4}$$

Hence, rather than assuming that a fixed dose-influence matrix D relates a fluence map x to a predictable dose distribution d, the dose distribution that will finally be delivered to the patient is unknown and may be given by any distribution d^k. How the scenario dose distributions d^k are calculated depends on the uncertainty under consideration and is further described in sections (4-6).

3.3. Formal approaches to robust planning

The next question is how the set of scenario dose distributions d^k can be incorporated into treatment planning. Each dose distribution d^k corresponds to an objective function value $f^k = f(d^k)$, which serves as a measure of treatment plan quality for error scenario k. Intuitively speaking, a treatment plan that is both good and robust yields a dose distribution d^k, which is good for all or the majority of error scenarios that may occur. There are different paradigms to translate this notion into mathematical terms. Broadly, these approaches can be categorized as follows:

1. The *stochastic programming approach* optimizes the expected plan quality.
2. The *minimax approach* optimizes plan quality for the worst error considered.

The stochastic programming approach and the minimax approach can be seen as extreme cases. In reality, one may be interested in controlling plan quality in between worst-case and average. Therefore, a third category of intermediate approaches should be considered. In this section, we formally define the different approaches; applications to specific uncertainties in IMRT and IMPT are discussed in sections (4-6).

3.3.1. Stochastic programming: In the stochastic programming, each error scenario is associated with an importance weight p_k. The approach then minimizes the expected value of the objective function:

$$\underset{x}{\text{minimize}} \quad \sum_k p_k f(d^k(x)) \tag{5}$$

The scenario weights p_k are often interpreted as the probability that error scenario k occurs. Hence, the stochastic programming approach minimizes the objective function

evaluated for all error scenarios, while more weight can be given to scenarios that are likely to occur and lower weight to extreme scenarios that are unlikely. In this review, we will refer to the parameters p_k as probabilities, however, it is worth noting that the application of equation 5 does not depend on such a probabilistic interpretation. The parameters p_k can simply be interpreted as weighting factors that indicate how important it is to achieve good plan quality for error scenario k. The stochastic programming approach is sometimes referred to as *probabilistic planning*, however, the term *probabilistic* is more broadly used also for other approaches that assign probability distributions to error scenarios. Stochastic programming has been widely applied in both IMRT [6, 7, 8, 9, 10] and IMPT [11, 12].

3.3.2. Minimax optimization The minimax approach aims at obtaining the treatment plan that is as good as possible for the worst error scenario that is considered:

$$\underset{x}{\text{minimize}} \quad \max_{k} \left[f(d^k(x)) \right] \tag{6}$$

Here, the maximum of the objective function over the error scenarios k is taken, which is minimized with respect to the incident fluence. In this case, no importance weights p_k are defined and the treatment plan only depends on the set of error scenarios. The approach is also referred to as *worst-case approach*. Minimax optimization has mostly been investigated in IMPT planning [13].

3.3.3. Minimax stochastic programming: The probabilistic approach and the minimax approach are related in the sense that a specific set of scenario weights p_k in the probabilistic approach yields the solution to the minimax formulation, namely such scenario weights that assign high p_k to the most unfavorable scenarios. The two methods can be interpreted as special cases of the *minimax stochastic programming* problem [14], which is defined as

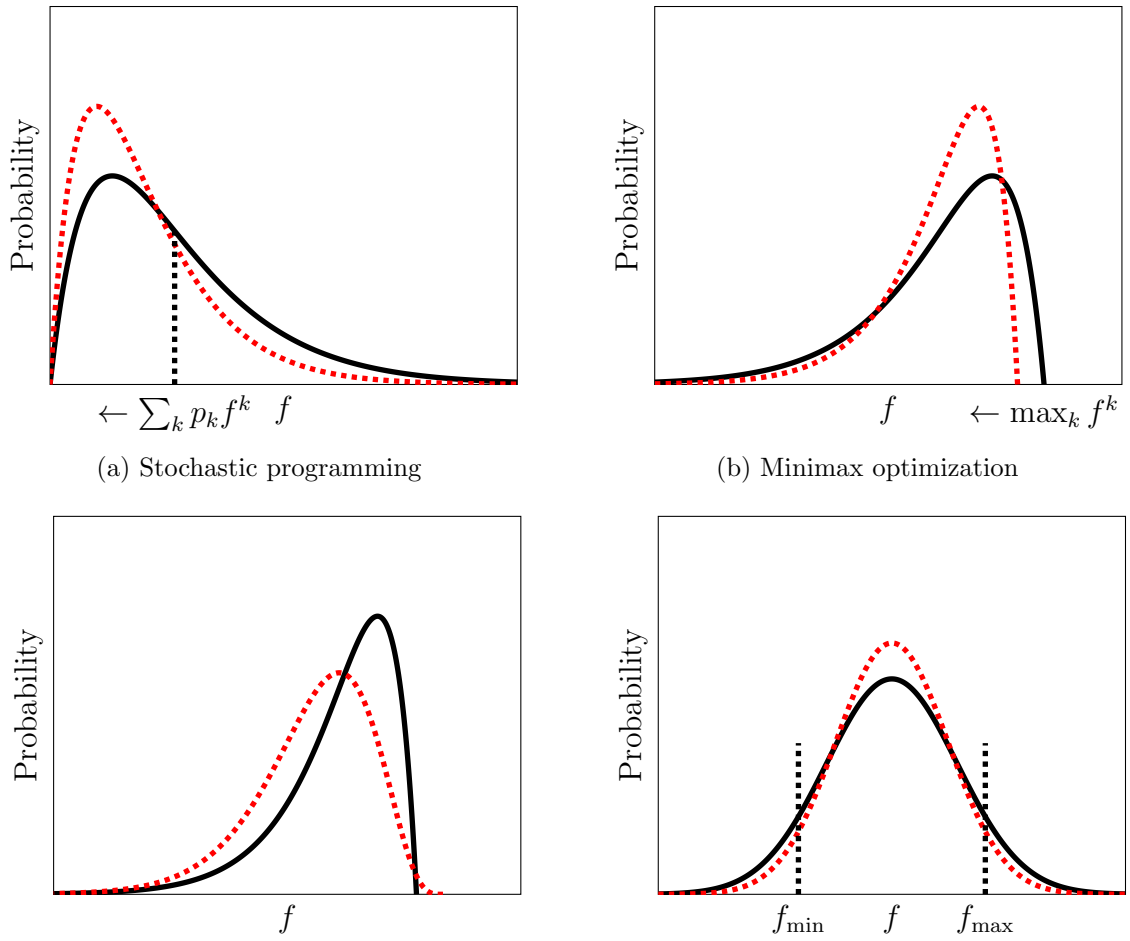
$$\underset{x}{\text{minimize}} \quad \max_{p \in \mathcal{P}} \quad \sum p_k f(d^k(x)), \tag{7}$$

where,

$$\mathcal{P} = \{p : 0 \leq p_k \leq \rho, \ \sum_k p_k = 1\} \tag{8}$$

is the uncertainty set for the scenario probabilities p_k. In words, this problem optimizes the expected value of the objective function for the most unfavorable probability distribution p_k over its uncertainty set. The parameter ρ controls the uncertainty level. For $\rho = 1$ every probability distribution is allowed, and consequently the minimax stochastic programming problem (9) is equivalent to the worst-case optimization in equation (6). For the choice $\rho = 1/K$, where K is the number of scenarios, we obtain the probabilistic approach (5) where equal importance $p_k = 1/K$ is assigned to all error scenarios. By selecting $1/K \leq \rho \leq 1$ one can gradually transition between the stochastic programming and worst-case formulation. This approach is also referred to as *distributionally robust* approach.

For the parameter choice $\rho = 1/(\alpha K)$ where $1/K \leq \alpha \leq 1$, minimax stochastic programming (9) is equivalent to what is referred to as *conditional value at risk* (CVaR) optimization with parameter α [14, 15, 16]. In CVaR optimization, the average of the fraction α of the worst scenarios is minimized. For example, for $\alpha = 0.1$ one would optimize the average plan quality for the worst 10% of scenarios, neglecting the 90% of more favorable scenarios.



(a) Stochastic programming

(b) Minimax optimization

(c) Minimax stochastic optimization

(d) Maximizing chances of correct treatment

Figure 1. Schematic illustration of different robust planning approaches. Uncertainty and a large number of error scenarios lead to a probability distribution over the objective function value, which serves as a plan quality indicator. The figures sketch the probability distribution (on the vertical axis) over objective function values (on the horizontal axis). Red dotted lines indicate an improved distribution, compared to black distributions, that the respected method is striving for. See Section 3.3.4 for discussion of (a-c) and Section 3.4.2 for discussion of (d).

3.3.4. Graphical illustration Let us interpret the scenario weights p_k as probabilities for the error scenario k to occur. If a large number of scenarios is considered, this results in a probability distribution over dose distributions d^k and consequently a probability

distribution over objective function values $f^k = f(d^k)$. Figure 1 schematically illustrates the above approaches. Figure 1a assumes the stochastic programming approach for an objective function $f \geq 0$ such as an NTCP function or the quadratic objective function to achieve a homogeneous target dose. The stochastic programming approach (equation 5) minimizes the mean of the distribution of objective function values, but not necessarily the tail towards large values. The treatment plan may allow for large objective function values for individual error scenarios, possibly corresponding large errors that are unlikely to occur. In contrast, the minimax approach (equation 6) only optimizes the maximum value of the distribution as illustrated in figure 1b. This usually requires that the set of error scenarios is truncated towards large errors. In addition, the minimax approach does not per se aim at improving the average objective value. The CVaR approach (figure 1c) represents an intermediate approach. It can be interpreted as a relaxation of the minimax method: unfavorable scenarios are emphasized without focusing purely on the worst case. Thereby, a small number of scenarios can have higher objective values for the benefit of better plan quality for most other scenarios.

3.4. Variations of these approaches

3.4.1. Robust constraints: Above, approaches for incorporating uncertainty in the objective function were considered. In addition, a treatment plan optimization problem may have constraints on the dose distribution. The most obvious approach for robustifying constraints is to enforce that the constraint is fulfilled for all error scenarios. This has mostly been investigated in proton therapy with a relatively small number of error scenarios, but also in the context of breathing motion [17] (see section 6.2).

3.4.2. Maximizing the probability of target coverage and OAR sparing: The probabilistic interpretation in figure 1 gives rise to variations of the approaches described above. One may want to achieve that a DVH criterion for target coverage is fulfilled for the majority of patients, for example, that in 95% of the scenarios, 95% of the target volume receices the prescribed dose. Similar to that, one may want to maximize the probability that a planning criterion is fulfilled. This is schematically illustrated in figure 1d. Here, it is assumed that f represents a dosimetric plan quality indicator such as an EUD or a DVH criterion. In this case, the goal is to minimize the cumulative probability that the value of f falls outside of the desired range, i.e. the probability that $f^k > f_{max}$, the probability that $f^k < f_{min}$, or both. For example, if f is the EUD in an OAR, one may want to minimize the cumulative probability that the EUD is larger than the maximum allowed EUD f_{max}. Hence the objective function to be minimized becomes

$$\underset{x}{\text{minimize}} \quad \sum_k p_k H(f(d^k(x)) - f_{max}) \tag{9}$$

where H denotes the heaviside step function, i.e. $H(f(d^k(x)) - f_{max}) = 1$ if $f(d^k(x)) > f_{max}$ and zero otherwise. Such approaches were investigated for interfraction motion in

IMRT by Sobotta et al [18] and Gordon et al [19].

3.4.3. Variations of the minimax approach There are several variations of the minimax approach in equation (6). In most practical situations, the objective function $f(d) = \sum_s w_s f_s(d)$ is a sum of objectives f_s for individual structures s weighted with importance factors w_s. In this case, it is possible to consider the maximum over scenarios for every structure-based objective individually rather than the maximum over the composite objective function:

$$\underset{x}{\text{minimize}} \quad \sum_s w_s \max_k \left[f_s(d^k(x)) \right] \tag{10}$$

This has been referred to as the *objective-wise worst case*, in contrast to the minimax approach in equation (6) which has been called the *composite worst case*. The objective-wise worst case approach has advantages in multi-criteria optimization and was investigated in that context [20]. Often, the objective function $f_s(d) = \sum_{i \in I_s} f_i(d_i)$ can further be written as a sum of contributions f_i from individual voxels i contained in structure s. In this case one can consider the maximum over the scenarios for each voxel separately, which leads to the *voxel-wise worst case* method:

$$\underset{x}{\text{minimize}} \quad \sum_s w_s \sum_{i \in I_s} \max_k \left[f_i(d_i^k(x)) \right] \tag{11}$$

This method can be interpreted as optimization of the worst-case dose distribution. To see this, one can consider the piece-wise quadratic objective function for overdosing of an OAR:

$$f(d) = \sum_i (d_i - d^{max})_+^2 \tag{12}$$

The contribution of a voxel i is determined by the maximum dose the voxel may receive for any scenario, i.e.

$$\max_k \left[\left(d_i^k - d^{max} \right)_+^2 \right] = \left(\max_k \left[d_i^k \right] - d^{max} \right)_+^2 \tag{13}$$

Hence, in this case, treatment plan optimization corresponds to evaluating the objective function for the worst-case dose distribution, which is defined, voxel-by-voxel, as the maximum dose delivered for any scenario. Similarly, a piece-wise quadratic objective for target underdosage can be considered, in which case the worst-case dose distribution corresponds to the minimum dose in each voxel. This approach was predominantly applied in robust IMPT planning [21, 22, 23].

3.4.4. Other approaches Chu et al. [24] presented another robust planning approach, whose starting point is that the dose and its uncertainty can be characterized by the expected dose and its variance. The probability that a voxel i in the CTV receives a dose higher than a prescribed minimum dose $d^{\min}$ depends on both the expected dose (which must be high enough) and the variance (which must be small enough). Assuming

$\sum_k p_k = 1$, the mean and variance of the dose in voxel i can be estimated based on the scenarios as

$$\mathbb{E}(d_i) = \sum_k p_k d_i^k \tag{14}$$

$$\mathbb{V}(d_i) = \sum_k p_k (d_i^k)^2 - (\mathbb{E}(d_i))^2 \tag{15}$$

where the expected dose is a linear function of the fluence x and the variance is a convex quadratic function of x. To ensure target coverage under uncertainty with high probability, Chu et al. propose a constraint on every voxel in the CTV such that the expected dose minus a multiple of the standard deviation exceeds the prescribed dose, i.e.

$$\mathbb{E}(d_i) - \delta\sqrt{\mathbb{V}(d_i)} \geq d^{\min} \tag{16}$$

Under the assumption that the dose in voxel i follows a Gaussian distribution (which is approximately the case for random errors but generally not for systematic errors) the parameter δ can be calculated based on the cumulative distribution function of the Gaussian and the accepted probability of underdosing. For example, limiting the probability of underdosing to 5% requires $\delta = 1.64$. The constraint (16) can be written as

$$\frac{d^{\min} - \mathbb{E}(d_i)}{\sqrt{\mathbb{V}(d_i)}} \leq \delta \tag{17}$$

which represents a *second order cone constraint*, which from an optimization perspective is almost as computationally efficient as a purely linear constraint. A similar constraint can be constructed for voxels in OARs and a maximum dose threshold $d^{\max}$.

Xie [25] presented a method that considers the expeced value and the variance of a general plan quality indicator rather than the dose in a voxel. Assume that f_s is a plan quality indicator associated with a structure s which has a desired value f_s^{pres}. Xie suggests to minimize

$$(\mathbb{E}(f_s) - f_s^{pres})^2 + \mathbb{V}(f_s) \tag{18}$$

where $\mathbb{E}(f_s)$ and $\mathbb{V}(f_s)$ are expected value and variance of the plan quality indicator, which are estimated from the scenarios analogously to equations (14) and (15). In the special case that f_s is the dose in a single voxel, this method is equivalent to the stochastic programming approach using the quadratic objective function as discussed in section 4.2. The method is suggested within a prioritized optimization framework to trade-off plan robustness against other plan quality measures.

3.5. General considerations and choice of method:

Among the different methods described above, no general statement can be made which method is superior. Research on handling interfraction motion in IMRT has largely

focused on the stochastic programming approach and other methods with probabilistic interpretation described in section 3.4.2. In the context of range and setup errors in IMPT, both stochastic programming as well as different versions of the worst-case method have been studied extensively. Any robust planning to handle specific types of uncertainty has to address several problems:

1. Modeling of the uncertainty. First, the uncertainty to be accounted for is to be modeled mathematically. This is straightforward for some types of errors, e.g. setup errors, which can be modeled as a rigid shift of the patient with respect to the isocenter. However, in other situations, e.g. breathing motion with variations in the breathing pattern, it is not immediately clear how to model the uncertainty.
2. Choosing an adequate method, i.e. formulating the robust optimization problem in a meaningful and computationally tractable way.
3. Developing ways to solve the optimization problem efficiently. This includes ways to calculate or approximate the dose distribution d^k for a given error scenario.

The above issues are interrelated. For example, whether a probabilistic or minimax approach is taken typically impacts what model of the uncertainty is suited. The worst case approach typically requires that the underlying error is truncated. For example, a treatment planner would set the maximum setup error to be accounted for. In the probabilistic approach instead, a setup error can be modeled via a Gaussian distribution, containing large errors with low weights p_k. Also, the need to devise efficient optimization algorithms influences the formulation of the problem as well as the model of the uncertainty.

4. Setup errors and inter-fraction organ motion in IMRT

In this section we review applications of robust planning for handling of setup errors and inter-fraction organ motion in IMRT. In this context, it is customary to separate the errors into a systematic component which is constant over all treatment fractions, and a random component which varies daily. A systematic error is typically introduced during treatment planning; an example is an extreme position of the prostate in the planning CT, which differs from the mean position of the prostate. The random error is related to daily patient anatomy and setup variations during fractionated treatments.

In section 4.1 we briefly discuss modeling of setup errors and inter-fraction motion. In section 4.2 we illustrate the stochastic programming approach applied to a stylized phantom, which provides insight into the qualitative features of this approach with respect to the handling of systematic and random errors. In the remaining part, we review the application to handling inter-fraction prostate motion, which has been the focus of most research. In section 4.7 we discuss computational aspects to facilitate robust planning at acceptable calculation times.

4.1. Modeling inter-fraction motion uncertainty

A setup error is usually understood as a rigid shift of the patient with respect to the isocenter. Most works in IMRT model setup errors using Gaussian distributions in 3 dimensions. Inter-fraction motion, for example of the prostate, is generally more complex to model. However, to first approximation, inter-fraction motion is often modeled as a setup error. This is appropriate if inter-fraction motion is mostly a rigid translation of the tumor inside the patient. This model can be improved by adding rotations. However, such an approach does not model deformations of the tumor and the surrounding anatomy. The most widely used approach for more accurate modeling of inter-fraction motion is principal component analysis (PCA) [26]. PCA identifies the dominant modes of deformable motion of the target and the surrounding anatomy. It yields a parmeterized model of the motion in which a plausible anatomical scenario is given by the mean position plus a linear combination of a set of eigenmodes multiplied by a scaling coefficient. In the case of prostate, this includes modeling expansion of rectum and bladder together with the resulting translation and rotation of the prostate in the sagittal plane. PCA modeling has a wide range of applications, including prostate dosimetric evaluation [27] and optimization [28] based on virtual treatment course simulation, coverage probability estimation [29], adaptive radiotherapy [30], and deformation modeling for lung [31]. Specifically for the deformable interfraction motion of prostate cancers, there have been several PCA models developed so far. Some models statistically model the surface deformations of two to three ROIs only [32, 33, 34]. Some other models model the entire 3D pelvic anatomy and therefore they are of more value to image guided adaptive radiotherapy applications. These models are either patient-specific or population based [35]. In general, the patient-specific models require an initial image data collection period to fully characterize the patient-specific motions. At the early phase of treatment where the patient data are very limited, the population-based model is more advantageous. However, the potential risk is that a population model may not benefit every patient if any unusual deformation is involved.

4.2. Qualitative features of the stochastic programming approach

The concept of stochastic programming can be illustrated by considering the quadratic objective function as an example:

$$f(d) = \sum_i (d_i - d^{pres})^2 \tag{19}$$

where d^{pres} is the prescribed homogeneous target dose and the sum runs over the voxels i in the target. Following the stochastic programming approach, the expected value of the quadratic objective function can be written as

$$\sum_k p_k f(d^k) = \sum_k p_k \sum_i \left(d_i^k - d^{pres}\right)^2 \tag{20}$$

$$= \sum_i (\mathbb{E}(d_i) - d^{pres})^2 \tag{21}$$

$$+\sum_i \left(\sum_k p_k \left(d_i^k - \mathbb{E}(d_i) \right)^2 \right) \tag{22}$$

where $\mathbb{E}(d_i)$ is the expected dose in voxel i as in equation (14). This has an intuitive interpretation. The first term (21) is the quadratic difference between the expected dose and the prescribed dose. Hence, minimizing this term will ensure that the dose will, on average, be close to the prescribed dose. The second term (22) represents the variance of the dose in each voxel. Minimizing this term will ensure that the dose d_i^k realized in a scenario k is close to the expected dose.

We now illustrate the result of stochastic programming using the quadratic objective function for a systematic Gaussian setup error in a one-dimensional phantom. The error scenarios correspond to shifts of the tumor up to ±10 mm. The probabilities p_k are chosen from a Gaussian distribution with 3 mm standard deviation. For this illustrative example, a simplified dose-influence matrix is assumed in which each beamlet j corresponds to a Gaussian beam profile with 3 mm standard deviation to model the penumbra. Setup errors are modeled as a shift of the voxel grid relative to the beamlet grid. Treatment plan optimization minimizes the expected value of the quadratic objective function, where the prescribed dose d^{pres} is set to 1 for the tumor and zero for the adjacent normal tissue.

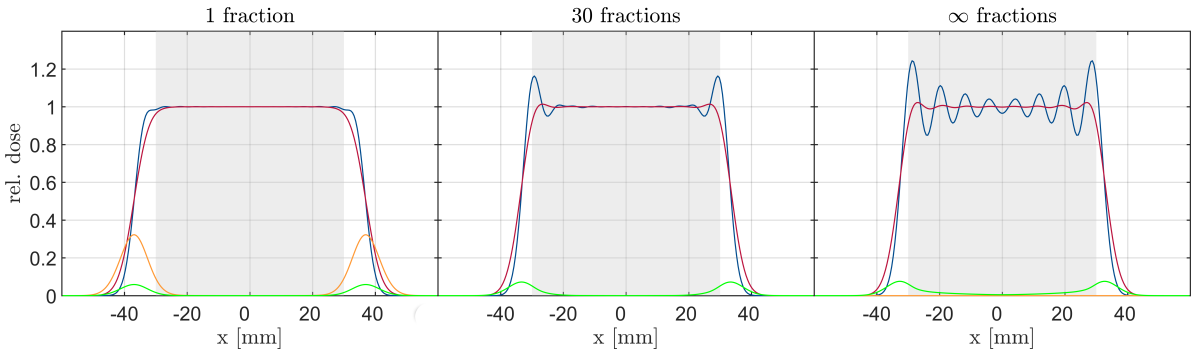


Figure 2. Illustration of stochastic programming for handling random setup errors in a one-dimensional phantom. The three panels correspond to a different number of fractions assumed for plan optimization: 1 fraction (left), 30 fractions (middle), and infinitely many fractions (right). We consider a 60 mm wide target volume in lateral direction, 36 beamlets spaced 2 mm apart corresponding to Gaussian beam profiles with 3 mm standard deviation, and a Gaussian setup uncertainty with 3 mm standard deviation. The nominal dose profile is shown in blue, the expected value of the dose is shown in red, and the standard deviation of the dose considering 30 fractions is shown in green. For comparison, the graphs for 1 and ∞ fractions also includes the standard deviation for the respective fractionation scheme used for optimization in orange.

A treatment plan which only considers a systematic error (which is equivalent to only considering random errors for one fraction) yields a dose distribution which is comparable to a PTV type treatment plan using conventional optimization (figure 2 left).

The incorporation of uncertainty into the optimization process induces an automatic expansion of the nominal dose (blue) around the CTV so that the expected dose (red) yields adequate coverage under uncertainty. Sir et al [36] investigated in more detail for systematic setup errors how the shape of the dose fall-off at the edge of the target volume depends on the type of objective function.

Next, we consider the case of random setup errors in a fractionated treatment with T fractions. An error scenario k now corresponds to a set of T setup errors $\{k_1, \dots k_T\}$. The total dose in scenario k is given by

$$d_i^k = \frac{1}{T} \sum_{t=1}^{T} d_i^{k_t} \tag{23}$$

The probability of an error scenario is given by the product of Gaussian probabilities for each fraction. In this case, the goal is to deliver a cumulative dose close to the prescription after the entire course of T fractions, whereas the dose in an individual fraction is allowed to vary. This leads to qualitatively different treatment plans as illustrated in figure 2 (middle). The nominal dose, i.e. the dose distribution delivered to a static geometry features *horns* as described in section 2. The treatment plan reduces the dose delivered to regions where both tumor and normal tissue can be located. As a consequence, the edge of the tumor may be underdosed in some fractions when the setup errors are large. However, the horns deliver doses higher than the prescription dose in some fractions, which compensates for fractions in which parts of the tumor is underdosed.

The height of the horns depends on the number of fractions. For a large number of fractions, the horns are more pronounced as more averaging will occur over the course of treatment. In the case that only a single fraction is delivered, a random error is equivalent to a systematic error in which no averaging occurs. In this case the horns disappear (figure 2 left). Mathematically, the number of fractions changes the relative weighting of the two terms in the objective function (21). The expected value of the dose is independent of the number of fractions while the variance decreases with the number of fractions as $1/T$ [37].

For the hypothetical case that infinitely many fractions are delivered, the expected value of the dose distribution is realized. Hence, the uncertainty in the dose distribution vanishes. The dose distribution under the influence of random errors is given by a convolution of the nominal dose distribution with the probability density function of the random error. Treatment planning can be performed analogously to conventional treatment planning except that the objective function is evaluated for the expected dose rather than the nominal dose. However, this approach emphasizes the horns (figure 2 right), which leads to dose uncertainty if the plan is delivered in a finite number of fractions. Mathematically, this solution can be interpreted as a deconvolution of a step

function (the desired ideal dose fall-off at the edge of the target volume) with the Gaussian probability distribution.

Figure 2 illustrates two main characteristics of the probabilistic approach when applied to setup errors. First, the approach allows for automated extension of the irradiated region around the target volume without explicitly defing a PTV. Second, random errors lead to qualitatively different plans featuring horns. These properties of the probabilistic approach have been demonstrated by several authors for stylized phantoms. Unkelbach et al [37, 38] considered a 2-dimensional rotation therapy model in conjunction with the expected value of the quadratic objective function. Earlier, Lind et al [39] and Löf et al [40] considered 1-dimensional phantoms together with TCP based objective functions. Recently, Witte et al [41] studied an asymmetric 2-dimensional model in which an OAR is located on one side of the tumor. The authors investigate the shape of the dose distribution that optimally balances tumor coverage and OAR sparing in the context of TCP as well as traditional objective functions. The authors also observed that, in the case that the residual random errors are small and are incorporated along with systematic errors, the tendency to generate horns is reduced.

4.3. Stochastic programming applied to inter-fraction prostate motion

The stochastic programming approach has been demonstrated for realistic patient geometries. The majority of these works considered the handling of inter-fraction motion of the prostate in IMRT [6, 42, 7, 9] but also head & neck cancer was considered as application [10, 43].

4.3.1. Handling of random errors: The paper by Unkelbach et al [6] and Maleike et al [42] apply stochastic programming with a quadratic objective function, which is demonstrated for a 1-dimensional phantom in section 4.2, to inter-fractional prostate motion. The analysis focuses on the handling of random errors. Qualitatively, the same effects can be observed as in figure 2. This includes the presence of horns for standard fractionated treatments in the static dose distribution that would be delivered to a static geometry, especially at the boundary of prostate and rectum. The works by Birkner et al [44] and McShan et al [45] investigate the handling of random errors by performing treatment plan optimization based on the expected value of the dose, which represents an approximation of the quadratic programming approach as described in section 4.2.

In principle, horns represent a mechanism to achieve steeper dose gradients at the edge of the target compared to conventional PTV based treatment plans. However, today the clinical desirability of dose horns is questionable: averaging out the inhomogeneities in the static dose distribution relies on the random errror being actually present during fractionated treatment. This may be in conflict with common thinking in practice, where the goal is to robustify a treatment plan against potential

random errors while simultaneously using all available means to avoid random errors. In addition, the work in [6] used relatively large random errors that are unrealistic in the era of image guidance. By most researchers, the concept of horns to sharpen the dose gradient at the edge of the target in the presence of random errors is therefore not considered a promising approach.

4.3.2. Stochastic programming using physical dose objectives: Stochastic programming was more extensively evaluated and compared to PTV based planning by Bohoslavsky et al [9] for prostate cancer and by Fontanarosa et al [10] for head & neck cancer. Bohoslavsky et al [9] developed an implementation of stochastic programming for interfraction motion as a plugin to a research version of the Pinnacle treatment planning system. The works optimize the expected value of objective functions typically used in clinical treatment planning, quadratic penalty functions and EUD objectives. A main finding of these works is that, using these tradiational objective functions, stochastic programming for handling systematic errors yields treatment plans that are qualitatively similar to PTV based plans. This is consistent with other publications demonstrating stochastic programming for systematic errors using quadratic penalty functions [6]. However, when OARs are located close to the CTV, stochastic programming may be used to redistribute dose away from OARs to less critical normal tissues, such that the dose to the OAR is lowered while the CTV remains covered under the majority of error scenarios.

In the context of dose painting by numbers, Witte et al [16] describe a modification of the stochastic programming approach towards conditional value at risk optimization. Instead of calculating the weighted average of the objective function over all scenarios, the summation occurs only over a subset of the scenarios. At each iteration of a gradient based optimization method, the scenarios are ranked according to their objective value. The summation is performed only over the better 90% of scenarios while neglecting the 10% of scenarios with the highest objective values. The method was applied to a quadratic penalty function with the goal of delivering at least the prescription dose to the CTV in 90% of the scenarios. In that sense, the method has similarities to the method by Gordon et al [19] described in section 4.5.

4.3.3. Optimizing expected TCP and NTCP: Stochastic programming using typical dose based objective functions, such as quadratic penalty functions, can automate the expansion of the irradiated region around the CTV. However, this does still not necessarily determine the optimal trade-off between target coverage and normal tissue sparing. In principle, stochastic programming is very natural in the context of TCP and NTCP based objective functions. A TCP model yields a value of tumor control probability for a given dose distribution. Taking the expectation over an uncertain dose distribution d^k,

$$\overline{TCP} = \sum p_k \mathrm{TCP}(d^k) \tag{24}$$

can be interpreted as the overall probability of tumor control taking into account geometrical uncertainty. Formally, this corresponds to marginalization over uncertain parameters in a probability distribution. Thereby, probabilistic planning could, at least in theory, find the optimal trade-off between target coverage and OAR sparing. The probabilistic approach applied to inter-fraction motion of the prostate using TCP based objective functions has been investigated by Witte et al [7]. One difficulty in evaluating the benefit of this approach is that using TCP based objectives usually results in treatment plans that are different from those used in clinical practice - due to the choice of the objective function rather than the way uncertainty is handled.

4.4. Maximizing the probability of adequate treatment

A variant of the concept outlined in section 3.4.2 was implemented in [18, 46]. In brief, the optimization problem was formulated as follows: minimize the cumulative probability that a dosimetric plan quality indicator for the target is worse than a given limit, while keeping the cumulative probability below a guaranteed maximum that dosimetric plan quality indicators for organs at risk are above a given limit. In other words, for each plan quality indicator, the objective or constraint is specified via a minimum or maximum limit and a maximum cumulative probability that this limit is exceeded. Conventional constraints on static plan quality indicators could be added.

Naturally, the crux of such a problem formulation lies in the estimation of the tails of the probability distributions of each plan quality indicator, which are usually sampled sparsely and are hence cost functions with a caveat. Sobotta et al suggest to approximate the cumulative probabilities of the tails by Chebyshev's inequalty, which reduces the problem to computing the mean and variance of the quality indicators' probability distributions. Still, for the treatment of inter-fractional uncertainties not satisfying the static dose cloud approximation, the computational burden of sampling the uncertainty space to estimate mean and variance can nevertheless become overwhelming. Sobotta et al further suggested to replace the direct evaluation of the plan quality indicators by a substitute patient model and demonstrated the utility of a Gaussian Process for this [47]. Other methods of machine learning may be suitable alternatives.

4.5. Probabilistic optimization of DVH objectives

A related approach to optimize plan quality for the majority of patients has been suggested by Gordon et al [19]. The work is motivated by the widespread use of DVH criteria for treatment plan evaluation, for example, that 95% of the target volume should receive the prescribed dose. An intuitive extension of this criteria in the context of setup uncertainty is to request that a DVH criterion is fulfilled for a given percentile of scenarios. This leads to the concept of *percentile dose volume histograms* (pDVH). In this approach, $d_{v,q}$ denotes the dose that is exceeded in the percentage volume v in

q percent of the scenarios. For example, a treatment planning goal can be to obtain a plan that delivers the prescription dose to 95% of the target in 90% of the patients, i.e. we would like $d_{95,90}$ to exceed the prescription dose.

A heuristic to obtain such a treatment plan has been suggested by Gordon et al [19]. At every iteration during treatment plan optimization, the current treatment plan is evaluated by sampling a large number of systematic setup errors and evaluating the dose distribution within the static dose cloud approximation. Based on that, $d_{v,q}$ is determined. Subsequently, an objective function is introduced that aims to increase $d_{v,q}$ to the prescription dose. To that end, a rim structure surrounding the CTV is introduced. This rim structures is analogous to a PTV, however, in contrast to traditional planning, the method does not aim to deliver the prescribed dose to all of the PTV but only to the degree necessary to achieve the desired $d_{v,q}$. In traditional planning, quadratic penalty functions are used as a heuristic to satisfy DVH objectives. Assume that the goal is to deliver d^{pres} to v% of the PTV. Further assume that the current plan does not fulfill this and that only a lower dose $d_v < d^{pres}$ is exceeded in v% of the PTV. Then, a quadratic penalty is introduced for all voxels in the PTV that receive a dose between d_v and d^{pres}:

$$f(d) = \sum_{i \in S} (d_i - d^{pres})_+^2 \tag{25}$$

where $S = \{i \in \text{PTV} \,|d_v < d_i < d^{pres}\}$. Hence, the quadratic penalty is applied to those voxels that are underdosed the least. This method can be modified to percentile DVH objectives by changing the set of voxels to $S = \{i \in \text{CTV} + \text{rim} \,|d_{v,q} < d_i < d^{pres}\}$. Here, the CTV+rim contains the voxels that play a role in achieving CTV coverage. Intuitively, it is clear that voxels close to the CTV are more important to ensure CTV coverage under the influence of errors, compared to voxels further away from the CTV edge. This can be incorporated by introducing voxel-dependent penalty factors in equation (25) such that voxels close to the CTV are weighted more than voxels further away.

With the same goal in mind [48] presented an extension to the work of [19] where they transferred the DVH-based coverage objectives into coverage constraints. Thereby they suggest a robust planning process that implements probabilistic constraints to avoid probabilistic criteria being traded in against competing objectives during optimization.

4.6. Optimization based on probability of tumor and organ presence

Baum et al [49] suggested a practical and computationally efficient method for handling systematic errors due to setup or inter-fraction organ motion. The method can be derived from the stochastic programming approach (equation 5) by a shift of perspective [50]. Robust planning methods as introduced in section 3 evaluate the dose in the patient's own coordinate system, where the dose becomes a random variable. Instead,

the approach by Baum et al considers the dose distribution in the treatment room coordinate system, which is constant within the validity of the static dose cloud assumption (section 4.7.1). Setup errors and inter-fraction motion can be modeled as a change in the segmentation of the planning CT into CTV and OARs, i.e. an error changes the set of voxels that belong to the CTV or an OAR. For each voxel i, the sum of all scenario probabilities p_k for a scenario k where i is occupied by a given structure corresponds to the probability that the voxel belongs to that structure, which Baum et al defined to be the *coverage probability* q_i. For any voxel-separable objective, these probabilities can be used as voxel-specific weighting factors in the objective function. For example, a quadratic penalty function for the CTV becomes

$$f(d) = \sum_i q_i \left(d_i - d^{pres}\right)^2 \quad . \tag{26}$$

This short derivation shows that the method essentially optimizes the mean objective function, averaged over all considered systematic errors. Optimization of this objective aims at delivering the prescribed dose to all voxels that may be occupied by the CTV or organ. However, if the probability is low, these voxels are weighted less. Hence, treatment plan optimization will preferentially lower the dose to voxels of the CTV if in conflict with other normal tissue objectives. The method has the advantage of being computationally efficient, adding a mere weight to every voxel of a classic static PTV or PRV patient model. Under certain conditions, namely validity of the static dose cloud approximation and voxel-separable objectives, this approach is mathematically equivalent to the stochastic programming approach for handling systematic errors. This has been shown for the quadratic objective function in [50]. While equation (26) describes the concept for achieving coverage of the CTV, the method is equally applicable to OAR objectives. In this case, q_i is the probability of voxel i being occupied by a certain OAR.

4.7. Computational considerations

4.7.1. Static dose cloud approximation: Application of robust planning techniques requires the evaluation of the dose distribution d^k for all scenarios under consideration. It would be possible to invoke the dose calculation algorithm several times and calculate a dose-influence matrix D^k for each error scenario. This would however be computationally and memory-wise expensive. Therefore, the dose distribution for an error scenario k is often approximated based on the nominal dose distribution. In photon therapy, approximate dose calculation of a voxel is based on the assumption that the effect of setup errors or inter-fraction motion can be approximated as a shift of the voxel relative to the nominal dose distribution. Let $\mathcal{D}(r)$ denote the nominal dose distribution as a function of the position r in 3-dimensional space, so that $d_i = \mathcal{D}(r_i)$ is the dose at voxel i whose center is located at position r_i. For an error scenario k, corresponding to setup error Δr^k, the dose in voxel i is approximated as $d_i^k = \mathcal{D}(r_i - \Delta r^k)$. The underlying assumption is that a setup error or a change of the patient geometry leaves the dose

distribution $\mathcal{D}$ in space unaffected, which has also been called the *static dose cloud approximation*. This is usually an acceptable approximation in photon therapy [51]. For proton therapy, this approximation breaks down and improvements are discussed in section 5.5.

4.7.2. Dose blurring for handling random errors: Many works approximate the effect of random errors via a convolution of the static dose cloud $\mathcal{D}$ with the probability distribution of a setup error, which is also referred to as dose blurring. In other words, this means that dose distribution resulting from random errors in T fractions as in equation (23) is replaced by the expected value of the dose and thereby becomes deterministic. This is valid for a treatment with infinitely many fractions and can be an acceptable approximation for standard fractionated treatments. However, it breaks down when the number of fractions is small. A related but different approximation strategy does not perform its random error blurring on the 3D dose distribution, but rather on the 2D fluence, projecting the errors in the patient's coordinate frame onto the planes perpendicular to each beam direction [52]. This fluence convolution approach may better handle those situations in which heterogeneous densities or air-tissue interfaces render the static dose cloud approximation invalid. While dose blurring is a good strategy for treatment plan evaluation, it should be noted that treatment plan optimization based on blurred dose distributions leads to the horns discussed in section 4.2, which may not be desired. Moore et al [53] demonstrated optimization incorporating fluence-convolution and found reduced OAR doses compared with margin-based plans.

4.7.3. Computational burden of robustness evaluation: Several approaches use sampling techniques to estimate the probability distribution over objective function values at each iteration [19, 18]. When dose distributions for errors are approximated by the static dose cloud approximation, it is interesting to note that the computation time needed to evaluate the probability distribution over objective function values is substantially less than one may intuitively expect. This is because the computation time is dominated by calculating the nominal dose distribution from the beamlet intensities. Using a typical beamlet size of 5 mm, an IMRT or VMAT plan may contain in the order of 10'000 beamlets. Hence, calculating the nominal dose to a single voxel requires 10'000 multiplications and additions. However, the nominal dose distribution needs to be calculated only once at each iteration during treatment plan optimization. Even if 10'000 dose distributions for errors are sampled subsequently, the computation time remains at the same order of magnitude [9].

5. Systematic range and setup errors in IMPT

Proton therapy has to deal with all the uncertainties encountered in photon therapy. However, there are two main additional challenges in proton therapy:

1. Range uncertainty. The finite range of protons is the main advantage of protons over photons, however, the exact position of the Bragg peak inside the patient is uncertain. To a large extent, range uncertainty arises from the problem that the planning CT image is an imperfect input for proton dose calculation. Imaging artifacts may corrupt the calculation of radiological depth, and the conversion of Hounsfield numbers to relative proton stopping power is not exact. In addition, pencil beam dose calculation algorithms widely used to reduce computation time compromise dose calculation accuracy in heterogeneous tissue compared to Monte Carlo algorithms.
2. The impact of setup errors and organ motion on the dose distribution tends to be more detrimental in proton therapy compared to photon therapy. This is in parts because changes in the geometry may lead to misalignment of tissue heterogeneities in the beam entrance path and thereby cause changes in the range or degradation of the Bragg peak. Therefore, setup errors do not simply lead to a shift of the dose distribution in the patient but may severely degrade the dose distribution.

Practitioners have always been aware of range uncertainties and proton-specific strategies to uncertainty handling were developed. These methods are more diverse and go beyond the PTV concept used in photon therapy. This includes the choice of beam directions that avoid areas of large tissue heterogeneities in the entrance path and avoid placing the distal field edge in front of an OAR. For the passive scattering technique, range uncertainty was addressed by increasing range and modulation of a spread-out bragg peak; setup uncertainty is addressed by widening the aperture; and compensator smearing is applied to account for misalignment of heterogeneities in the beam entrance path. Additional methods include the use of multiple patch field combinations to mitigate potential dose errors at the patch line. Interestingly, treatment planning and plan evaluation for passively scattered proton therapy was usually not based on a PTV concept. It has been suggested early on to evaluate treatment plans using multiple errors scenarios rather than evaluating PTV coverage [54]. In contrast to passively scattered proton therapy, treatment planning for pencil beam scanning proton therapy is nowadays based on mathematical optimization techniques. In that context, the PTV concept was applied in proton therapy, however, its inadequacy as the only means of uncertainty handling has been recognized. One of the main additional heuristics to achieve robustness is the single field uniform dose (SFUD) concept. As the name suggest, each individual beam direction delivers a uniform dose to a PTV, which typically reduces the sensitivity to range and setup errors. However, the SFUD technique sacrifices some of IMPT's potential for optimal sparing of normal tissues, especially for complex shaped target volumes that wrap around OARs. Over the past years, robust optimization methods for IMPT have been developed and are now implemented in several commercial treatment planning systems.

5.1. Limitations of the PTV concept in IMPT

The limitations of a PTV concept for handling range uncertainty are illustrated in figures 3a and 4a. The figures show an ependymoma patient treated with 3 fields. Gaussian pencil beams with an initial beam width of approximately 5 mm sigma have been used, corresponding to the latest generation of proton machines. Treatment planning aims at delivering a dose between 54 and 57 Gy(RBE) to the PTV using quadratic penalty functions. The PTV consists of a 2 mm isotropic expansion of the CTV. Additional objectives are quadratic penalty functions to limit the maximum dose in the brainstem to 54 Gy and for achieving conformity, minimization of the gEUD in the brainstem, and minimization of the mean dose in normal brain.

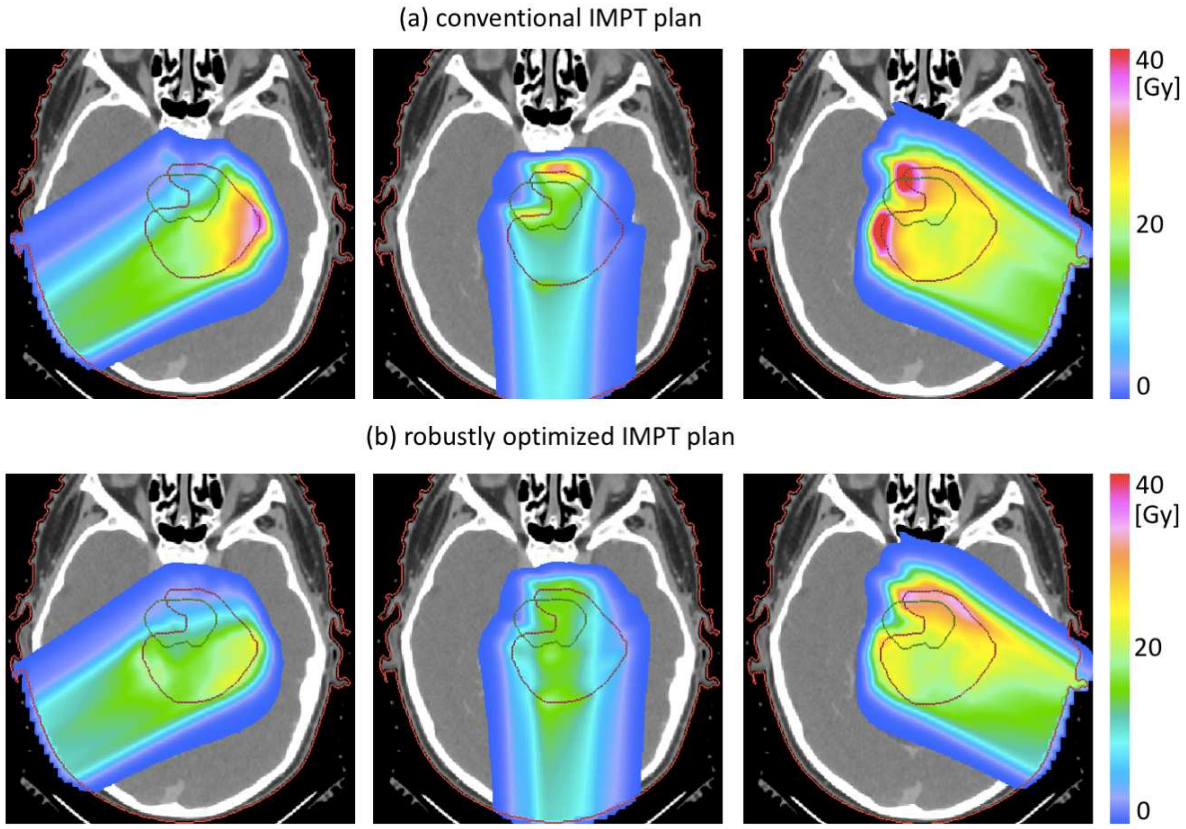


Figure 3. Three-beam IMPT plans for an ependymoma patient. (a) conventional PTV based treatment plan; (b) IMPT plan optimized for range uncertainty using the stochastic programming approach. Shown are the dose contributions of the three beams. The cumulative dose is shown in figure 4.

Figure 3a shows the contributions of the 3 fields. IMPT optimization tends to yield highly inhomogeneous dose contributions of individual fields. This is especially true if treatment planning aims at minimizing dose to healthy tissues. In this example, minimizing the mean dose in the normal brain leads to a preferential use of bragg peaks located at the distal edge of the target volume because these fields deliver dose to

the tumor for free while traversing the target volume. Figure 4a shows the cumulative dose distribution (left) together with the dose degradation observed for range overshoot (middle) and undershoot (right). A range error leads to a misalignment of the dose contribution of the three fields. In the case of a range overshoot (i.e. the range is larger than expected), the three dose contributions are shifted apart, leading to cold spots in the CTV. A range undershoot causes the dose contributions to be shifted closer together. This leads to increased doses in parts of the target volume, which may be undesirable in locations where the CTV overlaps with the brainstem. In both cases, range errors lead to inhomogeneous dose distributions within the target volume, a problem that can not be solved by PTV margins alone without additional heuristics such as SFUD.

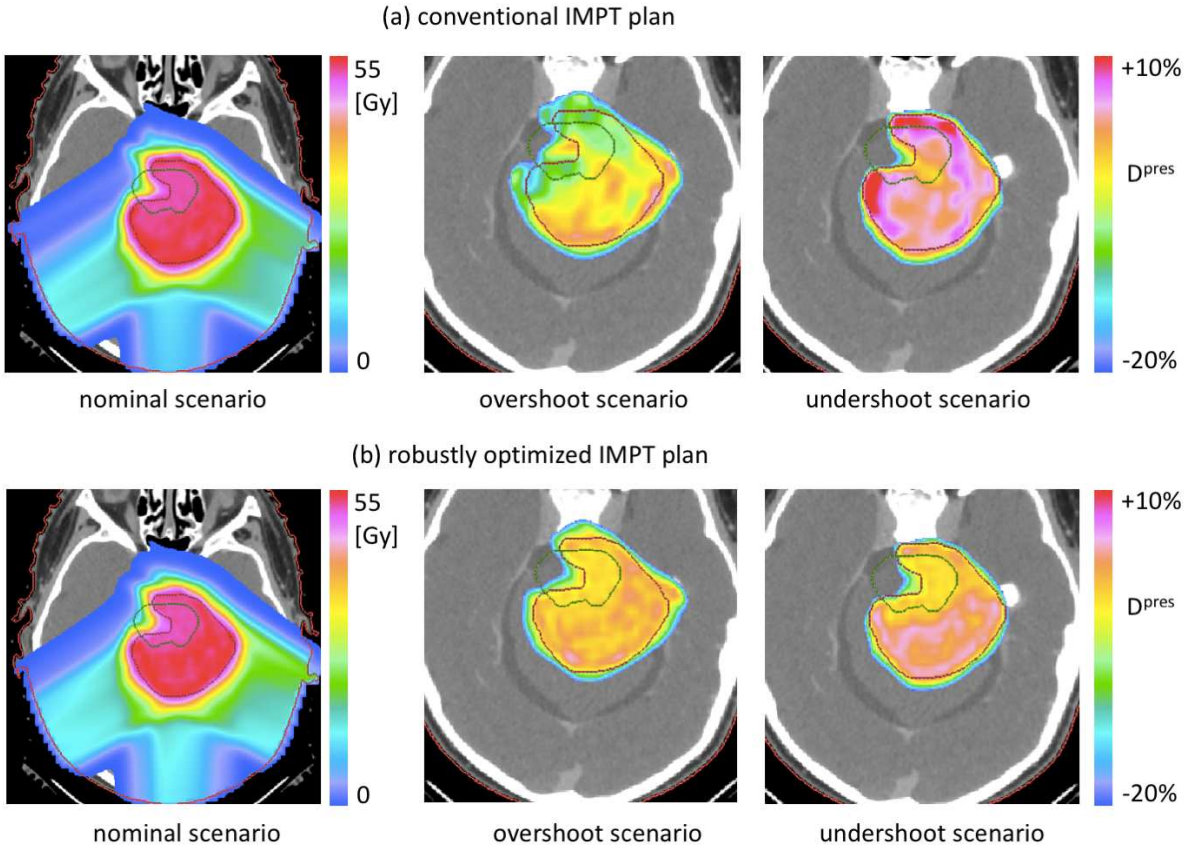


Figure 4. Robustness analysis of the IMPT plans shown in the figure 3. (left) nominal dose distribution; (middle) range overshoot; (left) range undershoot. Range over- and undershoot is modeled by down- and upscaling of the CT Hounsfield numbers by 4.6%.

Similarly, setup errors may lead to substantial degradation of the dose distribution rather than a simple shift of the dose as is approximately the case for photons [55, 12]. This has two reasons. First, a setup error has a different impact on each beam. Therefore, a setup error leads to misalignment of dose contributions similar to range uncertainties as illustrated above. This effect may occur even in homogeneous tissue. Second, setup errors may lead to misalignment of tissue heterogeneities. This is further

discussed in section 5.5. The degree of dose degradation depends on the amount of dose modulation in the dose contributions of individual beams and tends to be more severe the steeper dose gradients in these dose contributions are [56].

5.2. Qualitative features of robust planning

Important qualitative features of robust planning are illustrated in figures 3b and 4b. The stochastic programming approach is applied to robustify the IMPT plan of the ependymoma patient to range uncertainties. Uncertainty is modeled via 3 scenarios: the nominal scenario with a scenario weight of 0.5, and one scenario for range over- and undershoot with a weight of 0.25 each. Range errors are modeled by down- and upscaling of the CT hounsfield units by 4.6%. Features of robust planning for range uncertainty include:

- Dose gradients in beam direction in the dose contributions of individual beams are reduced. As a consequence, shifting these dose contributions in beam direction within the patient has a reduced impact on their cumulative dose distribution.
- The region proximal and distal to the CTV is irradiated to cover the CTV for range over and undershoot. In contrast to an isotropic PTV margin, the exansion of the irradiated region is created in a beam direction specific manner and depending on the dose contribution of the beam.

These features have been discussed in many early publications on robust planning [11, 21, 13]. Adding error scenarios for setup errors further modifes the treatment plan by expanding the irradiated region lateral to the CTV and by reducing dose gradients in the dose contributions of individual fields perpendicular to the beam direction [21].

5.3. Summary of robust IMPT planning publications

For IMPT planning, a large variety of methods have been studied. Table 1 provides an overview of publications focusing on those publications that introduce novel methods. The table summarizes the method used , the type of uncertainty accounted for, and the tumor site considered. So far, most works consider systematic range and setup errors. In recent years, a significant number of publications appeared that evaluate previously published methods or commercial implementations for various treatment sites [57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 15, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83].

- Unkelbach et al [11] published one of the first papers on robust IMPT planning. The work demonstrates stochastic programming as well as a voxel-wise worst case method for handling range uncertainty. The methods are demonstrated for a two-dimensional horseshoe shaped phantom and the qualitative features of robust plans are discussed.
- Pflugfelder et al [21] suggested the method of treatment plan optimization based on the worst-case dose distribution, corresponding to the voxel-wise worst case

Paper	Uncertainty	Method	Tumor site
Unkelbach [11]	range	SP, vWC	2D horseshoe phantom
Pflugfelder [21]	range + setup	vWC	paraspinal
Moravek [84]	dose alg. + organ motion	heuristic	lung
Unkelbach [12]	range + setup	SP	paraspinal
Fredriksson [13]	range + setup	cWC	lung, paraspinal, prostate
Inaniwa [85]	range + setup	heuristic	cervix, 2D phantom
Inaniwa [86]	range + setup	heuristic	prostate
Cao [87]	range + setup	vWC	prostate, skull base
Chen [20]	range + setup	oWC	chordoma, skull base
Fredriksson [88]	range + setup	SP, cWC, MSP	2D horseshoe phantom
Liu [89]	range + setup	vWC	prostate, skull base
Liu [22]	range + setup	vWC	lung, prostate, skull base
Bangert [90]	range + setup	SP	2D horseshoe phantom
Liu [91]	range + setup	vWC	head&neck
Petit [92]	range	vWC	liver, lung
Fredriksson [93]	setup	oWC	lung, prostate
Liu [94]	range + setup	vWC	base-of-skull
Fredriksson [95]	setup	cWC, oWC, vWC	prostate
Liu [96]	range + setup	vWC	lung
Liu [97]	range + setup + organ motion	vWC	lung
Bokrantz [98]	range + setup + organ motion	SP	lung, prostate, 3D phantom
An [15]	range + setup	SP	lung, prostate, head&neck
Wahl [99, 100]	range + setup	SP	prostate, paraspinal, intracranial

Table 1. Overview of publications on robust IMPT planning, summarizing the method used, the uncertainties addressed and the treatment site used for demonstration. SP = stochastic programming; cWC, vWC, oWC = minimax optimization in the flavors composite, voxel-wise, and objective-wise worst case; MSP = minmax stochastic programming.

method as described in section 3.4.3. The uncertainty model includes range and setup errors, modeled via 9 scenarios. The method is demonstrated for a horseshoe shaped tumor surrounding the spinal cord. The method was then further studied by Liu et al [89, 22] and evaluated for several tumor sites including head & neck [91], base-of-skull [94], and lung [96, 97].

- Fredriksson et al [13] introduced minimax optimization to IMPT optimization and demonstrated the method to a lung, prostate, and paraspinal tumor.
- Chen et al [20] investigated the objective-wise worst case method in the context

of multi-criteria optimization (MCO). In this case, the worst case is evaluated for each individual objective, which makes the method suited for MCO.

- Fredriksson [88] describes the minimax stochastic programming model that can continuously interpolate between the minimax and stochastic programming approach. The three methods are compared for a for a two-dimensional horseshoe shaped phantom.
- Knowing the qualitative features of robust plans can be used to develop heuristics for robust treatment planning. Inaniwa et al [85, 86] add terms to the objective function that suppress in-field dose gradients and pencil beams that may deliver a high dose to an OAR, instead of performing scenario based robust optimization. This provides some of the benefits of robust optimization at significantly reduced computational cost.
- Bangert et al [90] devised an analytical probabilistic modeling framework bypassing sampling for stochastic programming. A fully Gaussian parameterization of the underlying dose calculation enables closed form computation and hence optimization of the expected value of the quadratic objective function.
- Bokrantz and Fredriksson [98] introduced a scenario-based method that is equivalent to geometric margins if the scenario doses are calculated using the static dose cloud approximation. If more accurate scenario doses are used, then the method provides a comparable level of robustness as the minimax and stochastic approaches while simultaneously avoiding some of their disadvantages.

5.4. Comparison of methods

For handling range and setup errors in IMPT, a large variety of robust planning methods has been studied. So far, there is no consensus that one particular method is generally superior. To first approximation, all methods provide the same fundamental advantage over margins, i.e. the use of scenario dose distributions d^k that provide a physically realistic model of the dosimetric effect of errors. Thereby, misalignments of tissue heterogeneities or dose contributions of individual fields are accounted for irrespective of whether stochastic programming, minimax optimization or an intermediate approach is taken. That stochastic programming and worst-case approaches can give qualitatively similar results is a finding originally presented by Unkelbach et al [11].

On a more detailed quantitative level, different methods may or may not yield different results depending on geometry, uncertainty, and planning objectives. The most extensive comparison of robust planning methods is provided by the publications of Fredriksson and Bokrantz [95, 101, 98], showing that some methods may yield undesirable results in specific situations. Figure 5 illustrates this for the case of setup errors. Figure 5a shows the planned dose for a prostate target without OARs in its vicinity. The depicted results were generated by optimization with respect to quadratic penalties on deviation from the prescription dose within the target and deviation from zero dose

elsewhere. The weight for healthy tissue sparing was set small enough as not to compromise target coverage. Systematic setup shifts were discretized into scenarios using a uniform grid with a step size of 1/3 cm. Shifts up to 1 cm in 3D were accounted for in the optimization, which resulted in 123 scenarios in total. The uncertainty for the probabilistic formulation was assumed to follow a Gaussian distribution with 5 mm sigma that was truncated at two standard deviations. Stochastic programming (left) and minimax optimization (right) yield almost identical results (figure 5a).

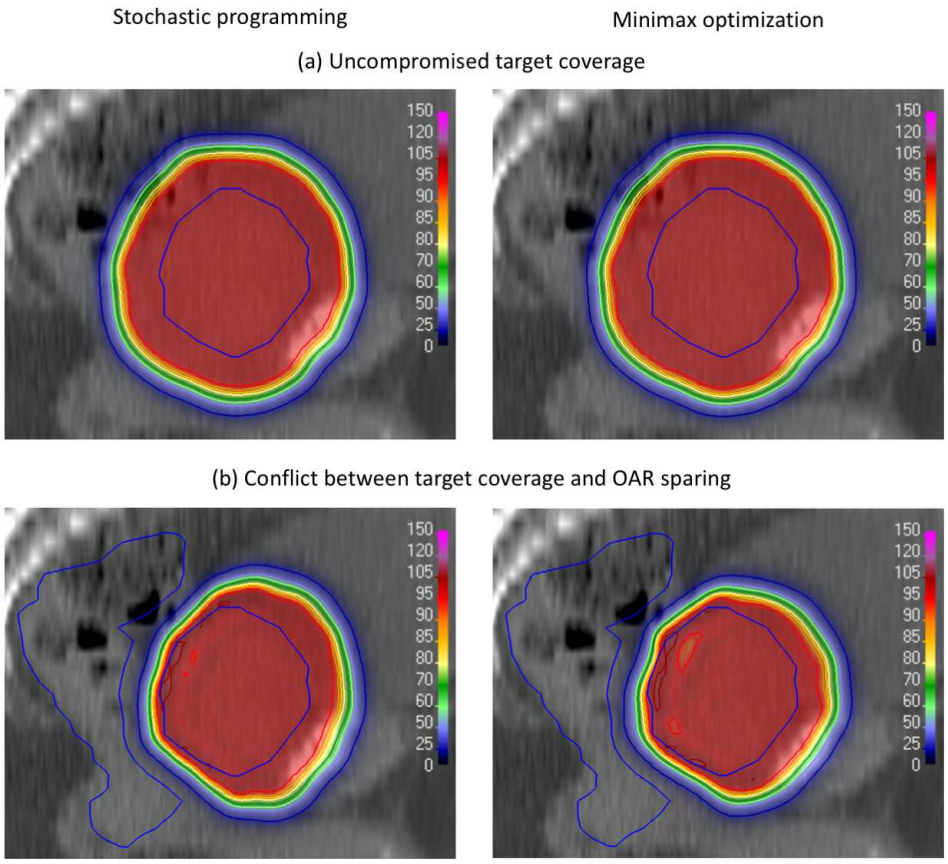


Figure 5. Comparison of stochastic programming (left) and minimax optimization (right) for a prostate cancer case. The figures show a sagittal slice. In (a) all normal tissue surrounding the target is weighted equally. In this case, both robust planning approaches yield almost identical treatment plans. In (b) dose to the rectum is penalized more such that target coverage is compromised for a setup error in posterior direction. In this case, the two approaches yield distinct results.

A clearer distinction can be made when the desired target dose must be compromised due to adjacent OARs. This may yield circumstances when the worst-case approach has undesired consequences. Figure 5b shows the same example as that shown in Figure 5a, except that the objective function f is augmented with a term that emphasizes sparing of the rectum. The weight for this term was set high enough to be in considerable conflict with target coverage. The depicted results show that the mini-

max method can unnecessarily neglect easy scenarios where target coverage need not be compromised (shifts along the inferior-superior axis) if a severe conflict between targets and OARs exist under other scenarios (shifts along the posterior-anterior axis). A similar example can be found in Fredriksson et al [95]. These examples illustrate that the worst-case approach is more sensitive to the definition of the uncertainty set than the probabilistic approach. To resolve conflicts between OAR sparing and target coverage, the minimax approach may require explicit selection of the scenarios against which to be robust.

The choice of robust planning method may also take the formulation of the objective function f into account. If the objective function is a probability measure, such as TCP or NTCP, the stochastic programming approach is more natural as described in section 4.3.3. Similarities with a PTV margin regarding how target coverage is traded against OAR sparing may be an argument in favor of the worst-case approach if treatment plan optimization is performed with respect to standard physical penalties such as those used in Figure 5a. Further examples that compare the stochastic programming and minimax approaches with respect to physical penalties can be found in Fredriksson et al [101] and Bokrantz et al [98].

5.5. Approximation of error dose distributions

To quantify the effects of possible errors, robust radiotherapy planning methods require the dose distributions d^k under multiple error scenarios. In photon therapy, these are often approximated based on the static dose cloud approximation as described in section 4.7.1. In proton therapy, this approximation is usually insufficient. This section outlines some of the advanced methods for calculating or approximating scenario dose distributions d^k. The results of the methods are illustrated for scenario dose calculation for a systematic setup error on a lung case that has been planned with a 5 mm PTV rather than any robust treatment planning method. The nominal dose in a transversal slice for this case is shown in Figure 6a. Because of the heterogeneous density of the treatment site and the failure to use robust treatment planning, it is expected that the dose under the setup error will be deformed compared to the nominal dose, and no longer cover the target.

5.5.1. Separate dose-influence matrices The most accurate way of determining the effects of errors on the dose distribution is to perform an accurate dose calculation under each scenario. Separate dose-influence matrices are then stored for each scenario during the optimization. The dose under each scenario is as accurate as the nominal dose. The method is expensive in terms of memory and computation time, however, this could be addressed through high-performance computing. Figure 6b shows the accurately calculated dose under a setup shift of 5 mm to the patient's right (left side of the image). Because of the heterogeneous density of the site, the dose deforms as an

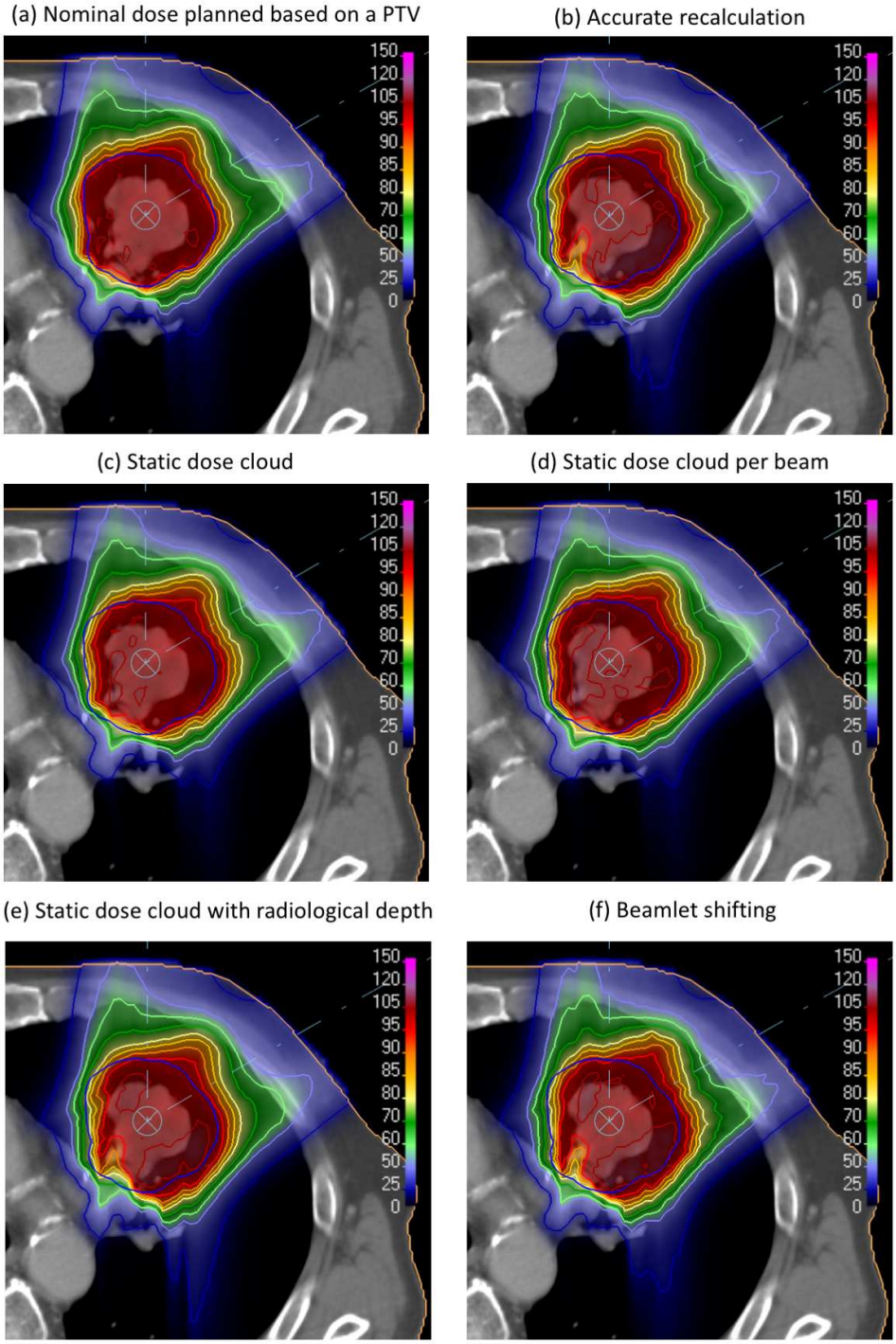


Figure 6. Comparison of different methods to approximate error dose distributions in IMPT for a 5 mm setup error to the patient's right (left side of the image). Changes of the radiological depth along the path of a proton pencil beam resulting from a setup error are approximated well only by a subset of methods.

effect of the error. In this example, this leads to a cold spot in the distal part of the target volume.

5.5.2. Voxel shifting methods: A computationally cheap way of calculating scenario doses is to perform transformations of the nominal dose distribution. In this case, the dose to a voxel for a given error is approximated based on the nominal dose to a different voxel. Here, three methods of doing so are presented.

1. The static dose cloud approximation: This method is described in Section 4.7.1 and has been extensively used for photons. Figure 6c shows the dose distribution according to the static dose cloud approximation under a setup shift of 5 mm to the patients right. Because the static dose cloud approximation does not deform the dose, it reflects the effect of the setup error poorly. Its $\gamma(3\%/3\text{ mm})$ pass rate for voxels with accurately calculated dose above 10% of its maximum was 88.7%.
2. The static dose cloud approximation per beam: The static dose cloud approximation can be improved based on the observation that a setup shift affects the dose contribution from different beam directions differently. For example, a setup shift along the beam direction does only marginally affect the dose delivered by this beam, but impacts the dose from other beam directions. Therefore, a setup error may, even in the absence of tissue heterogeneities, lead to dose degradation due to misalignment of the dose contributions from different beams. To account for this, the static dose cloud approximation can be applied to each beam separately by calculating an effective voxel shift taking into account the direction of the setup error, the beam direction, and the orientation of the patient's surface [12]. Figure 6d shows the dose distribution according to the static dose cloud approximation per beam under a setup shift of 5 mm to the patient's right. For the lung patient, the dose difference indicates that there is little benefit in using the static dose cloud approximation per beam as compared to the standard static dose cloud approximation. Its $\gamma(3\%/3\text{ mm})$ pass rate for voxels with accurately calculated dose above 10% of its maximum was 90.2%. This is because the dose degradation is dominated by the misalignment of tissue heterogeneities (which is not accounted for) rather than the misaligment of dose contributions from different fields.
3. Voxel shifting accounting for radiological depth: A further improved voxel shifting method takes the radiological depth into account during the voxel shifting. The beam dose to each voxel is not only shifted according to the setup error projected onto the plane perpendicular to the beam central axis, but is also shifted in the direction parallel to the beam to a point with the same radiological depth as the voxel had prior to the shifting [102]. Figure 6e shows the dose distribution according to the voxel shifting method taking radiological depths into account under a setup shift of 5 mm to the patients right. The approximated dose reflects the deformation that occurs due to the setup shift. Its $\gamma(3\%/3\text{ mm})$ pass rate for voxels with accurately calculated dose above 10% of its maximum was 99.4%.

5.5.3. Beamlet shifting: Beamlet shifting moves the approximation from the voxel domain to the fluence domain. This way, the dose calculation algorithm's ability to take

the density distribution of the patient into account is utilized also for the scenario dose calculation, but the computational effort is still much reduced compared to calculating full dose-influence matrices for each scenario. In the beamlet shifting method, a setup error is approximated as a shift of the spot weights (in a fixed spot grid) according to the setup error projected onto the plane perpendicular to the beam. The spot weights at the lateral edges of the spot grid are to be shifted to spots outside the spot grid, which necessitates the calculation of the dose of virtual spots in an extended spot grid, i.e. spots that are only used in the scenario dose calculation but are excluded from the plan. Moreover, if the spots are shifted to positions for which no dose-influence has been calculated, interpolation must be used. To improve on the interpolation, virtual spots can be calculated between the planned spot positions [12]. Figure 6f shows the dose distribution according to the beamlet shifting under a setup shift of 5 mm to the patient's right. Its $\gamma(3\%/3\text{ mm})$ pass rate for voxels with accurately calculated dose above 10% of its maximum was 99.8%. The differences between the doses arise because shifting of beamlets does not account for beam divergence. Thus, the greater the source axis distance, the smaller the approximation becomes.

5.5.4. Approximate dose calculation for range errors All approximation methods except the static dose cloud can handle range errors. The static dose cloud per beam would shift each beam dose longitudinally using geometric depth; static dose cloud with radiological depth would shift each beam dose longitudinally using radiological depth; beamlet shifting would shift the spot weights to other energy layers. For all methods, the distance shifted is a function of depth or energy.

5.6. Analytical probabilistic modeling

In section 3, approaches to robust planning are formulated using a discrete set of error scenarios. In some applications, the set of error scenarios has been small. For example, in IMPT robust planning models with 9 error scenarios have been studied, consisting of the nominal scenario, range overshoot, range undershoot, and 6 setup errors. In IMRT applications a much larger number of scenarios is typically considered to more accurately represent a Gaussian distribution. In any case, errors are discretized for numerical integration. Analytical probabilistic modeling [90] is an alternative approach for uncertainty quantification, which is primarily studied in the conjunction with stochastic programming. The main idea behind analytical probabilistic modeling is to use a functional parameterization of pencil beam dose distributions via Gaussian distributions. Using also Gaussian distributions for range and setup errors enables closed-form integration to directly compute the expected value and the standard deviation of the dose, and in some cases the objective function value. Consequently, the full continuous probability density describing uncertainty can be incorporated; it is not necessary to compute individual scenarios.

For proton therapy under range and setup uncertainty, analytical probabilistic modeling provides more consistent estimates of the expected value and the standard deviation of the dose at reduced computation times compared to sampling approaches [99]. This translates into computational advantages for stochastic programming using the quadratic objective function [99]. The computational advantages are particularly prominent in the context of fractionated radiotherapy as analytical probabilistic modeling allows for a consistent incorporation of random and systematic sources of uncertainty [100]. Recently, it has been shown that the analytical probabilistic modeling framework also generalizes to the non-linear computations of the relative biological effectiveness of carbon ions at the same computational complexity [103].

Analytical probabilistic modeling has the potential to enable novel approaches to uncertainty management based on an analytical definition and differentiation of probabilistic objectives and constraints. However, it does not easily generalize to non-pencil beam dose calculation algorithms providing higher accuracy and also the incorporation of uncertainties beyond patient setup and particle range (e. g. anatomical deformations) is an open question.

6. Respiratory motion

Commonly, motion compensation strategies for handling respiratory motion are divided into 3 types of approaches: 1) gating, where the treatment beam is turned off when the tumors is outside a defined region, 2) tracking, where motion of the treatment couch or the MLC leafs is used to compensate for motion in real time, and 3) safety margin approaches. The latter includes the internal target volume (ITV) approach, where the target volume is defined as the union of the target volumes in all respiratory phases obtained from a 4D CT. An alternative margin approach is the mid-ventilation concept where an appropriate margin is added to the target volume defined in the mid-ventilation phase. A fourth approach that incorporates the motion into treatment plan optimization is often forgotten. Such methods can broadly be divided into two groups:

1. Approaches that assume a probability density function (PDF) for the position of the target, which is incorporated into plan optimization. The motion PDF describes the relative amount of time that the tumor spends in each breathing phase. In this case a single treatment plan is created, which is delivered without any online adjustments to motion measured during treatment.
2. Approaches that assume that the motion is predictable or monitored in real time, and that the delivery of radiation can be synchronized with the motion.

Both approaches have sometimes been referred to as 4D optimization despite being rather different. Therefore, we avoid this term in this review. The second type of approach is difficult in terms of delivery. In the case of photon therapy, it could be considered an extension of MLC tracking. We briefly review these works in section 6.4,

however, we focus on the first type of approach that relates to robust planning more directly.

6.1. Treatment plan optimization based on a known motion PDF

The ITV approach for moving tumors, and the PTV approach in general, aim to deliver the prescribed dose to all regions were the tumor may be, regardless of how much time the tumor spends in each position. In the presence of motion, and when the total tumor dose is achieved by accumulating dose contributions from multiple geometric instances, the ITV approach is suboptimal in terms of normal tissue sparing. In particular, normal tissue dose can be reduced by delivering less dose to regions where the tumor is rarely. To ensure that this dose reduction does not compromise target coverage, higher doses should be delivered to regions largely occupied by the tumor.

To formalize this concept, assume that a lung tumor accumulates dose over different phases φ of the breathing cycle, and let w^φ be the nonnegative fraction of time spent in each phase. Each phase has an associated dose-influence matrix D^φ, whose calculation usually involves deformable image registration to map voxels in each breathing phase to their location in a reference phase. Let us assume that the total dose accumulated over a breathing cycle can be approximated as

$$d = \sum_\varphi w^\varphi D^\varphi x = \bar{D}x\,, \quad \sum_\varphi w^\varphi = 1 \tag{27}$$

where

$$\bar{D} = \sum_\varphi w^\varphi D^\varphi \tag{28}$$

is an effective dose-influence matrix. It is important to note that w^φ is not a scenario probability but the fraction of time spent in phase φ. Further, it is important to distinguish motion from uncertainty. Up to this point, no uncertainty is considered, i.e. it is assumed that the cumulative dose is given by blurring a nominal dose distribution with the known motion PDF. This is similar to the handling of random errors discussed in section 4 for an infinite number of fractions. Using a fixed effective dose-influence matrix $\bar{D}$ in IMRT optimization will create a treatment plan featuring horns, which is optimal for the assumed motion pattern characterized by w^φ.

This approach was studied by various authors. Söhn et al [104] used the probability density function (w^φ) to explicitly optimize the accumulated dose under respiratory motion for lung patients and showed that their plan generated similar results as gating. Zhang et al [105] concluded that the approach can achieve plans similar to those achieved by real-time target tracking. Watkins et al [106] and Lens et al [107] compared the motion PDF approach and ITV planning, and showed that it resulted in similar target coverage but significantly lower dose to surrounding healthy tissues. In the study by Watkins et al [106], target coverage and OAR sparing was mostly maintained when

the motion PDF differed from that assumed during optimization. The motion PDF approach was also studied in IMPT in combination with the voxel-wise worst-case method for handling range and setup uncertainty [97].

6.2. Robust planning for handling uncertainty in the motion PDF

A treatment that is optimized for a fixed motion PDF may degrade if the actual breathing pattern varies substantially from the assumed motion PDF w^{φ}. Robust planning methods can be used to robustify the plan against uncertainty in the breathing pattern. One approach to parameterize this uncertainty is to assume uncertainty in w^{φ}, i.e. uncertainty in the amount of time spent in each phase. This can be done by defining a set of possible breathing patterns w_k^{φ}, which translates into a set of effective dose-influence matrices $\bar{D}^k$. Subsequently, any of the robust planning concepts described in section 3.3 such as the minimax or stochastic programming approach can be applied. However, in this case it is also possible to consider a continuous uncertainty set $\mathcal{W}$, containing a set of realistic breathing PDFs.

This robust planning approach is a generalization of both the ITV approaches and optimization based on a fixed PDF. The fixed PDF approach can be recovered by simply setting the uncertainty set $\mathcal{W}$ to be a single vector equal to w^{φ} for all φ. At the other end of the spectrum, the largest set that $\mathcal{W}$ could be is the unit simplex: $\{w^{\varphi} \ : \ \sum_{\varphi} w^{\varphi} = 1, w^{\varphi} \geq 0, \forall \varphi\}$. This set models the situation where the breathing motion can be any possible breathing pattern, including ones where the patient spends 100% of the time at a single phase in the breathing cycle. This is representative of the ITV approach. An intermediate choice of $\mathcal{W}$ results in a solution that balances between healthy tissue sparing and target coverage under breathing motion uncertainty. Thus, the robust optimization approach generates a continuum of robustness that allows the decision maker to modulate the degree of conservatism when designing the treatment [108].

Several robust planning approaches for handling respiratory motion including uncertainty in breathing patterns have been investigated [17, 109, 8]. Unkelbach et al [109] investigated a stochastic programming approach for handling uncertain respiratory motion. The work first considered uncertainty in the motion PDF as well as breathing amplitude and baseline variations in a 1-dimensional phantom. This work was expanded by Heath et al [8] and demonstrated for a lung cancer patient. The latter work also provides a comparison of stochastic programming to the voxel-wise worst case method described in section 3.4.3. Chan et al [17] were the first to propose a robust optimization approach, which was demonstrated in a simplified 1-dimensional phantom. Later, this approach was generalized and demonstrated in a lung patient geometry, with the formal mathematical development of the continuum of robustness defined above [108].

This approach was then adapted to optimize the DVH tails in breast IMRT under respiratory motion uncertainty, where the key trade-off was between cardiac sparing and target coverage [110]. This approach demonstrated the potential to reduce the need for breath-hold techniques [111], without requiring much extra computational effort to solve the more challenging, tail DVH-based robust optimization problem [112].

One of the consistent findings from robust treatment planning approaches is the presence of horns in the dose distribution that would be delivered to a static geometry, which are designed such that edge-enhancements are washed out by the motion so that the prescribed cumulative dose is delivered to the target volume. For robust planning approaches, the horns are smoother and less pronounced compared to treatment planning based on a fixed breathing PDF. In fact, depending on the degree of uncertainty in the breathing PDF, the approach can interpolate between an ITV-like treatment plan and the fixed PDF situation.

In contrast to the case of random errors for setup and inter-fraction motion, horns may be an acceptable approach to handle motion in the case of breathing motion. In fact, in lung or liver SBRT highly non-uniform dose distributions are delivered in clinical practice, which show hot spots of up to 150% of the prescription dose inside the target volume. Although, these treatment plan may be motivated by other considerations, this also facilitates target coverage with smaller margins. This aspect has been investigated theoretically by Chan et al [113, 114], showing that horns can be optimal in dealing with motion. Vranvcic et al [115] provided experimental validation by delivering horn-based fluence maps on a linear accelerator. McCann et al [116] and Ahanj et al [117] showed that edge-enhanced intensity maps at inhale and smaller beam apertures during inhale can provide the same coverage as margins but potentially reduce the dose to healthy tissue for lung cancer.

6.3. Combining adaptation with robust optimization

Adaptive radiation therapy describes a broad paradigm of closed-loop decision making where parameters are updated as a treatment progresses to improve the quality of the final treatment [118]. Although not originally proposed for dealing with uncertain respiratory motion, the concept was adapted to this case to further improve the performance of previously proposed robust optimization models, which are limited by a fundamental trade-off: a larger uncertainty set results in better target coverage at the expense of increased normal tissue dose. To push this trade-off frontier forward, Chan et al [119] proposed an adaptive robust optimization approach where the uncertainty set is updated from fraction to fraction based on past observations of the patient's breathing pattern. They showed that simultaneous improvements in both target coverage and normal tissue sparing were possible, and provided theoretical justification to support the effectiveness of their approach. Subsequent research showed that the adaptive robust

approach also performed well when considering daily dose metrics [120] and breathing patterns that drifted over time [121].

6.4. Plan optimization assuming synchronization of tumor motion and delivery

Another approach to incorporate respiratory motion in treatment plan optimization consist in optimizing a separate fluence map for every respiratory phase. In this case, the cumulative dose distribution

$$d = \sum_{\varphi} D^{\varphi} x^{\varphi} \tag{29}$$

is given by the sum of doses $D^{\varphi}x^{\varphi}$ delivered in each phase, where x^{φ} is the fluence map delivered while the patient is in phase φ. The objective function is evaluated for the cumulative dose and minimized with respect to all fluence maps simultaneously. In principle, this approach can provide a treatment that improves even on the current approach to MLC or couch tracking. In the current approach to tracking, a treatment plan is optimized for one respiratory phase. During delivery using MLC tracking, the apertures of the treatment plan are shifted to compensate for target translation. The above approach can in principle improve on that. A treatment that allows for distinct fluence maps for each resiratory phase can treat different parts of the target volume primarily in the respiratory phase that is the most suited, e.g. when the target moves away from an OAR during inhale or exhale. In other words, this approach does not only mitigate motion, but may *exploit* motion to improve a plan over the static situation.

This approach has been investigated by Trofimov et al [122] and compared to other motion handling approaches. Nohadani et al [123] added constraints on the fluence maps to the treatment plan optimization problem to enforce similarity of fluence maps for neighboring respiratory phases. In these works, the approach of delivering separate fluence maps for each respiratory phase was investgated conceptually, however, the question how such treatment plans would be delivered efficiently was not addressed.

Obtaining a deliverable treatment plan requires an optimization method that is aware of the delivery process. In photon therapy, this is the case for direct aperture optimization. Assuming that both the motion of the tumor and the delivery of a treatment plan over time is known a priori, each aperture can be assigned to a particular breathing phase. In this case, a set of apertures is optimized based on their cumulative dose, assuming that each aperture is delivered during a known breathing phase. Such an approach was investigated for VMAT and for IMRT planning by several groups [124, 125, 126]. Similarly, in proton therapy a predetermined scan path can be considered such that each pencil beam spot can be assigned to a given breathing phase during plan optimization. The cumulative dose to a voxel i can then be written as

$$d_i = \sum_{\varphi} \sum_{j} z_{j\varphi} D_{ij}^{\varphi} x_j \tag{30}$$

where $z_{j\varphi}$ is a binary indicator that assigns pencil beam j to phase φ, i.e. $z_{j\varphi} = 1$ if pencil beam j is delivered during phase φ and zero otherwise. Such an approach was investigated by Bernatowicz et al [127] and others. In this approach, a single fluence map is optimized, but different beamlets are assigned to different phases. The resulting treatment plan can therefore not reach the theoretical optimum where a separate fluence map per phase is optimized, however, the approach would substantially improve on any margin approach. The review by Bert et al [128] presents a detailed review of respiratory motion management in proton therapy. So far, these approaches have assumed a perfect synchronization of tumor motion and treatment delivery and did not consider uncertainty in the delivered dose. In that sense, they represent a method to incorporate respiratory motion in treatment planning, but not a robust planning method to account for uncertainty in planning and delivery. Engwall et al. [129] applied robust optimization to account for uncertainties in breathing motion and delivery synchronization while optimizing a single proton spot scanning pattern to be delivered over the different phases. The consideration of multiple breathing motion scenarios resulted in reduced sensitivity to the interplay effect due to irregularities in the breathing motion for the considered patients.

7. Discussion

The fundamental limitations of the PTV concept in IMPT led to the first implementations of robust planning in commercial treatment planning systems. Thereby, robust planning has evolved from a research topic to a methodology used in clinical practice of proton therapy planning. RayStation (RaySearch Laboratories) supports robust optimization for photons, protons, and carbon ions based on the composite worst-case approach (equation 6). Robust optimization for protons is also available in Eclipse (Varian), which features an implementation of the voxelwise worst-case approach (equation 11) similar as described by Liu et al [89], and in Pinnacle (Philips Healthcare), which follows the probabilistic approach (equation 5) [130].

All three of the commercial implementations can take patient setup uncertainty and particle range uncertainty into consideration. The magnitude of the setup shifts to be accounted for is generally specified separately for left-right, anterior-posterior, and superior-inferior direction; the magnitude of particle range errors to be accounted for is specified in percent of the nominal range. RayStation can also handle organ motion by using multiple existing patient images, such as the phases of a 4D-CT, or by generating synthetic images that simulate organ motion. Robust optimization in RayStation was the first commercial implementation and has been evaluated for protons in several works [60, 61, 62, 63, 64, 65, 66, 67]. Similar studies were done for Eclipse [131, 132, 59]

The limitations of the PTV approach are less severe in IMRT compared to IMPT, and robust planning for systematic errors based on quadratic penalty functions yields

treatment plans that are often qualitatively similar to PTV based treatment plans. Situations where robust planning would have a fundamental advantage such as TCP/NTCP based optimization or dose painting [16], are not commonly done in practice so far. Perhaps therefore, the application of robust planning for IMRT has lagged behind that for IMPT, even though methods like stochastic programming were initially investigated for geometric uncertainty in IMRT and were only later applied to IMPT. Raystation is the only commercial planning system that supports the use of robust optimization for IMRT planning, which has been evaluated in a number of publications in recent years [133] including applications to breast [134, 135], lung [136, 137], and glioblastoma [138].

Even though IMRT planning systems do not commonly support robust optimization methods, for some applications, it is possible to mimick the nature of the robust solution, obtained from a research TPS, with a commercial one. For the case of boosting lymph nodes of cervix patients in a simultaneous integrated boost technique, the coverage probability approach by Baum [49] (section 4.6) was established and clinically validated [139, 140] by a transfer of dose plan features obtained from the experimental TPS Hyperion to Varian Eclipse RapidArc plans. The positive experience with this technique have led to planning goals for the EMBRACE II cervix cancer trial that derive from robust planning concepts, and not PTV concepts [141].

References

[1] J. C. Stroom, H. C. J. De Boer, H. Huizenga, and A. G. Visser. Inclusion of geometrical uncertainties in radiotherapy treatment planning by means of coverage probability. *International Journal of Radiation Oncology Biology Physics*, 43(4):905–919, 1999.

[2] M. van Herk, P. Remeijer, C. Rasch, and J. V. Lebesque. The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 47(4):1121–1135, 2000.

[3] M. Van Herk. Errors and margins in radiotherapy. In *Seminars in radiation oncology*, volume 14, pages 52–64. Elsevier, 2004.

[4] J. J. Gordon and J. V. Siebers. Evaluation of dosimetric margins in prostate IMRT treatment plans. *Medical physics*, 35(2):569–75, 2008.

[5] D. Bertsimas, O. Nohadani, and K. M. Teo. Nonconvex robust optimization for problems with constraints. *INFORMS journal on computing*, 22(1):44–58, 2010.

[6] J. Unkelbach and U. Oelfke. Incorporating organ movements in IMRT treatment planning for prostate cancer: Minimizing uncertainties in the inverse planning process. *Med. Phys.*, 32(8):2471–83, 2005.

[7] M. G. Witte, J. van der Geer, C. Schneider, J. V. Lebesque, M. Alber, and M. van Herk. IMRT optimization including random and systematic geometric errors based on the expectation of TCP and NTCP. *Med. Phys.*, 34(9):3544–3555, 2007.

[8] E. Heath, J. Unkelbach, and U. Oelfke. Incorporating uncertainties in respiratory motion into 4D treatment plan optimization. *Med. Phys.*, 36(7):3059–3071, 2009.

[9] R. Bohoslavsky, M. G. Witte, T. M. Janssen, and M. Van Herk. Probabilistic objective functions for margin-less IMRT planning. *Phys. Med. Biol.*, 58(11):3563, 2013.

[10] D. Fontanarosa, H. P. van der Laan, M. Witte, G. Shakirin, E. Roelofs, J. A. Langendijk, P. Lambin, and M. van Herk. An in silico comparison between margin-based and probabilistic

target-planning approaches in head and neck cancer patients. *Radiother. Oncol.*, 109(3):430–436, 2013.

[11] J. Unkelbach, T. C. Y. Chan, and T. Bortfeld. Accounting for range uncertainties in the optimization of intensity modulated proton therapy. *Phys. Med. Biol.*, 52(10):2755–2773, 2007.

[12] J. Unkelbach, T. Bortfeld, B. C. Martin, and M. Soukup. Reducing the sensitivity of IMPT treatment plans to setup errors and range uncertainties via probabilistic treatment planning. *Med. Phys.*, 36(1):149–163, 2009.

[13] A. Fredriksson, A. Forsgren, and B. Hårdemark. Minimax optimization for handling range and setup uncertainties in proton therapy. *Med. Phys.*, 38(3):1672–1684, 2011.

[14] A. Fredriksson. A characterization of robust radiation therapy treatment planning methods—from expected value to worst case optimization. *Med. Phys.*, 39(8):5169–5181, 2012.

[15] Y. An, J. Liang, S. E. Schild, M. Bues, and W. Liu. Robust treatment planning with conditional value at risk chance constraints in intensity-modulated proton therapy. *Med. Phys.*, 44(1):28–36, 2017.

[16] M. Witte, G. Shakirin, A. Houweling, H. Peulen, and M. Van Herk. Dealing with geometric uncertainties in dose painting by numbers: Introducing the ??vH. *Radiotherapy and Oncology*, 100(3):402–406, 2011.

[17] T. C. Chan, T. Bortfeld, and J. N. Tsitsiklis. A robust approach to IMRT optimization. *Phys. Med. Biol.*, 51(10):2567, 2006.

[18] B. Sobotta, M. Söhn, and M. Alber. Robust optimization based upon statistical theory. *Med. Phys.*, 37(8):4019–4028, 2010.

[19] J. Gordon, N. Sayah, E. Weiss, and J. Siebers. Coverage optimized planning: probabilistic treatment planning based on dose coverage histogram criteria. *Med. Phys.*, 37(2):550–563, 2010.

[20] W. Chen, J. Unkelbach, A. Trofimov, T. Madden, H. Kooy, T. Bortfeld, and D. Craft. Including robustness in multi-criteria optimization for intensity-modulated proton therapy. *Phys. Med. Biol.*, 57(3):591–608, 2012.

[21] D. Pflugfelder, J. J. Wilkens, and U. Oelfke. Worst case optimization: a method to account for uncertainties in the optimization of intensity modulated proton therapy. *Phys. Med. Biol.*, 53(53):1689–1700, 2008.

[22] W. Liu, X. Zhang, Y. Li, and R. Mohan. Robust optimization of intensity modulated proton therapy. *Med. Phys.*, 39(2):1079–1091, 2012.

[23] W. Liu, S. J. Frank, X. Li, Y. Li, P. C. Park, L. Dong, X. R. Zhu, and R. Mohan. Effectiveness of robust optimization in intensity-modulated proton therapy planning for head and neck cancers. *Med. Phys.*, 40:051711, 2013.

[24] M. Chu, Y. Zinchenko, S. G. Henderson, and M. B. Sharpe. Robust optimization for intensity modulated radiation therapy treatment planning under uncertainty. *Phys. Med. Biol.*, 50(23):5463, 2005.

[25] Y. Xie. *Applications of Nonlinear Optimization*. Ph.D. Thesis, Graduate School of Arts and Sciences, Washington University, St. Louis, USA, 2014. DOI: 10.17605/OSF.IO/EUMZJ, https://osf.io/eumzj/.

[26] M. B. Christopher. *Pattern recognition and machine learning*. Springer-Verlag New York, 2016.

[27] M. Söhn, B. Sobotta, and M. Alber. Dosimetric treatment course simulation based on a statistical model of deformable organ motion. *Phys. Med. Biol.*, 57(12):3693, 2012.

[28] H. Xu, D. J. Vile, M. Sharma, J. J. Gordon, and J. V. Siebers. Coverage-based treatment planning to accommodate deformable organ variations in prostate cancer treatment. *Med. Phys.*, 41(10), 2014.

[29] G. J. Price and C. J. Moore. A method to calculate coverage probability from uncertainties in radiotherapy via a statistical shape model. *Phys. Med. Biol.*, 52(7):1947, 2007.

[30] S. Thörnqvist, L. B. Hysing, A. G. Zolnay, M. Söhn, M. S. Hoogeman, L. P. Muren, and B. J. Heijmen. Adaptive radiotherapy in locally advanced prostate cancer using a statistical

deformable motion model. *Acta Oncologica*, 52(7):1423–1429, 2013.

[31] Q. Zhang, A. Pevsner, A. Hertanto, Y.-C. Hu, K. E. Rosenzweig, C. C. Ling, and G. S. Mageras. A patient-specific respiratory model of anatomical motion for radiation treatment planning. *Med. Phys.*, 34(12):4772–4781, 2007.

[32] M. Söhn, M. Birkner, D. Yan, and M. Alber. Modelling individual geometric variation based on dominant eigenmodes of organ deformation: implementation and evaluation. *Phys. Med. Biol.*, 50(24):5893, 2005.

[33] S. Thörnqvist, L. B. Hysing, A. G. Zolnay, M. Söhn, M. S. Hoogeman, L. P. Muren, L. Bentzen, and B. J. Heijmen. Treatment simulations with a statistical deformable motion model to evaluate margins for multiple targets in radiotherapy for high-risk prostate cancer. *Radiother. Oncol.*, 109(3):344–349, 2013.

[34] E. Budiarto, M. Keijzer, P. Storchi, M. Hoogeman, L. Bondar, T. Mutanga, H. de Boer, and A. Heemink. A population-based model to describe geometrical uncertainties in radiotherapy: applied to prostate cases. *Phys. Med. Biol.*, 56(4):1045, 2011.

[35] D. J. Vile. *Statistical modeling of interfractional tissue deformation and its application in radiation therapy planning.* Virginia Commonwealth University, 2015.

[36] M. Y. Sir, S. M. Pollock, M. A. Epelman, K. L. Lam, and R. K. Ten Haken. Ideal spatial radiotherapy dose distributions subject to positional uncertainties. *Phys. Med. Biol.*, 51(24):6329, 2006.

[37] J. Unkelbach and U. Oelfke. Inverse planning incorporating organ movements via probability distributions of voxel locations. *Radiother. Oncol.*, 73(Sup. 1):S347, 2004.

[38] J. Unkelbach and U. Oelfke. Incorporating organ movements in inverse planning: assessing dose uncertainties by Bayesian inference. *Phys. Med. Biol.*, 50:121–139, 2005.

[39] B. K. Lind, P. Källman, B. Sundelin, and A. Brahme. Optimal radiation beam profiles considering uncertainties in beam patient alignment. *Acta Oncologica*, 32(3):331–342, 1993.

[40] J. Löf, B. K. Lind, and A. Brahme. Optimal radiation beam profiles considering the stochastic process of patient positioning in fractionated radiation therapy. *Inverse Problems*, 11:1189–1209, 1995.

[41] M. G. Witte, J.-J. Sonke, J. Siebers, J. O. Deasy, and M. van Herk. Beyond the margin recipe: the probability of correct target dosage and tumor control in the presence of a dose limiting structure. *Physics in Medicine & Biology*, 62(19):7874–7888, 2017.

[42] D. Maleike, J. Unkelbach, and U. Oelfke. Simulation and visualization of dose uncertainties due to interfractional organ motion. *Phys. Med. Biol.*, 51(9):2237–2252, 2006.

[43] D. Fontanarosa, M. Witte, G. Meijer, G. Shakirin, J. Steenhuijsen, D. Schuring, M. van Herk, and P. Lambin. Probabilistic evaluation of target dose deterioration in dose painting by numbers for stage ii/iii lung cancer. *Practical radiation oncology*, 5(4):e375–e382, 2015.

[44] M. Birkner, D. Yan, M. Alber, J. Liang, and F. Nüsslin. Adapting inverse planning to patient and organ geometrical variation: algorithm and implementation. *Med. Phys.*, 30(10):2822–31, 2003.

[45] D. McShan, M. Kessler, K. Vineberg, and B. Fraass. Inverse plan optimization accounting for random geometric uncertainties with a multiple instance geometry approximation (miga). *Med. Phys.*, 33(5):1510–1521, 2006.

[46] B. Sobotta. *Optimization of the Robustness of Radiotherapy Against Stochastic Uncertainties.* PhD Thesis, 2011.

[47] B. Sobotta, M. Söhn, and M. Alber. Accelerated evaluation of the robustness of treatment plans against geometric uncertainties by gaussian processes. *Phys. Med. Biol.*, 57(8):8023–8039, 2012.

[48] H. Mescher, S. Ulrich, and M. Bangert. Coverage-based constraints for IMRT optimization. *Phys. Med. Biol.*, 62(18):N460, 2017.

[49] C. Baum, M. Alber, M. Birkner, and F. Nüsslin. Robust treatment planning for intensity modulated radiotherapy of prostate cancer based on coverage probabilities. *Radiother. Oncol.*, 78:27–35, 2006.

[50] J. Unkelbach and U. Oelfke. Relating two techniques for handling uncertainties in IMRT optimization. *Phys. Med. Biol.*, 51:N423–N427, 2006.

[51] M. Sharma, E. Weiss, and J. V. Siebers. Dose deformation-invariance in adaptive prostate radiation therapy: Implication for treatment simulations. *Radiotherapy and oncology*, 105(2):207–213, 2012.

[52] W. A. Beckham, P. J. Keall, and J. Siebers. A fluence-convolution method to calculate radiation therapy dose distributions that incorporate random set-up error. *Phys. Med. Biol.*, 47:3465–3473, 2002.

[53] J. a. Moore, J. J. Gordon, M. S. Anscher, and J. V. Siebers. Comparisons of treatment optimization directly incorporating random patient setup uncertainty with a margin-based approach. *Medical physics*, 36(9):3880–3890, 2009.

[54] M. Goitein. Calculation of the uncertainty in the dose delivered during radiation therapy. *Med. Phys.*, 12(5):608–612, 1985.

[55] A. J. Lomax. Intensity modulated proton therapy and its sensitivity to treatment uncertainties 2: the potential effects of inter-fraction and inter-field motions. *Phys. Med. Biol.*, 53(4):1043–1056, 2008.

[56] F. Albertini, E. B. Hug, and A. J. Lomax. Is it necessary to plan with safety margins for actively scanned proton therapy? *Phys. Med. Biol.*, 56(14):4399–4413, 2011.

[57] Y. Li, P. Niemela, L. Liao, S. Jiang, H. Li, F. Poenisch, X. R. Zhu, S. Siljamaki, R. Vanderstraeten, N. Sahoo, et al. Selective robust optimization: A new intensity-modulated proton therapy optimization strategy. *Med. Phys.*, 42(8):4840–4847, 2015.

[58] L. Liao, G. J. Lim, Y. Li, J. Yu, N. Sahoo, H. Li, M. Gillin, X. R. Zhu, A. Mahajan, S. J. Frank, et al. Robust optimization for intensity modulated proton therapy plans with multi-isocenter large fields. *International Journal of Particle Therapy*, 3(2):305–311, 2016.

[59] K. Stützer, A. Lin, M. Kirk, and L. Lin. Superiority in robustness of multifield optimization over single-field optimization for pencil-beam proton therapy for oropharynx carcinoma: An enhanced robustness analysis. *Int. J. Radiat. Oncol. Biol. Phys.*, 99(3):738–749, 2017.

[60] M. Stuschke, A. Kaiser, C. Pöttgen, W. Lübcke, and J. Farr. Potentials of robust intensity modulated scanning proton plans for locally advanced lung cancer in comparison to intensity modulated photon plans. *Radiother. Oncol.*, 104(1):45–51, 2012.

[61] M. Stuschke, A. Kaiser, J. Abu Jawad, C. Pöttgen, S. Levegrün, and J. Farr. Multi-scenario based robust intensity-modulated proton therapy (IMPT) plans can account for set-up errors more effectively in terms of normal tissue sparing than planning target volume (PTV) based intensity-modulated photon plans in the head and neck region. *Radiat Oncol*, 8:145, 2013.

[62] M. Stuschke, A. Kaiser, J. Abu Jawad, C. Pöttgen, S. Levegrün, and J. Farr. Re-irradiation of recurrent head and neck carcinomas: comparison of robust intensity modulated proton therapy treatment plans with helical tomotherapy. *Radiat Oncol*, 8:93, 2013.

[63] L. V. van Dijk, R. J. Steenbakkers, B. ten Haken, H. P. van der Laan, A. A. van 't Veld, J. A. Langendijk, and E. W. Korevaar. Robust intensity modulated proton therapy (IMPT) increases estimated clinical benefit in head and neck cancer patients. *PloS one*, 11(3):e0152477, 2016.

[64] T. Inoue, J. Widder, L. V. van Dijk, H. Takegawa, M. Koizumi, M. Takashina, K. Usui, C. Kurokawa, S. Sugimoto, A. I. Saito, et al. Limited impact of setup and range uncertainties, breathing motion, and interplay effects in robustly optimized intensity modulated proton therapy for stage III non-small cell lung cancer. *Int. J. Radiat. Oncol. Biol. Phys.*, 96(3):661–669, 2016.

[65] M. Cubillos-Mesías, M. Baumann, E. G. C. Troost, F. Lohaus, S. Löck, C. Richter, and K. Stützer. Impact of robust treatment planning on single- and multi-field optimized plans for proton beam therapy of unilateral head and neck target volumes. *Radiation Oncology*, 12(1):190, Nov 2017.

[66] S. Tang, L. Song, J. D. Sturgeon, and C. Chang. Robust planning for a patient treated in decubitus position with proton pencil beam scanning radiotherapy. *Cureus*, 9(9), 2017.

[67] A. Barragán, S. Differding, G. Janssens, J. A. Lee, and E. Sterpin. Feasibility and robustness

of dose painting by numbers in proton therapy with contour-driven plan optimization. *Med. Phys.*, 42(4):2006–2017, 2015.

[68] S. van de Water, F. Albertini, D. C. Weber, B. J. Heijmen, M. S. Hoogeman, and A. J. Lomax. Anatomical robust optimization to account for nasal cavity filling variation during intensity-modulated proton therapy: a comparison with conventional and adaptive planning strategies. *Phys. Med. Biol.*, 63(2):025020, 2018.

[69] S. Van Der Voort, S. van de Water, Z. Perkó, B. Heijmen, D. Lathouwers, and M. Hoogeman. Robustness recipes for minimax robust optimization in intensity modulated proton therapy for oropharyngeal cancer patients. *Int. J. Radiat. Oncol. Biol. Phys.*, 95(1):163–170, 2016.

[70] M. Zaghian, W. Cao, W. Liu, L. Kardar, S. Randeniya, R. Mohan, and G. Lim. Comparison of linear and nonlinear programming approaches for "worst case dose" and "minmax" robust optimization of intensity-modulated proton therapy dose distributions. *Journal of applied clinical medical physics*, 18(2):15–25, 2017.

[71] J. Yu, X. Zhang, L. Liao, H. Li, R. Zhu, P. C. Park, N. Sahoo, M. Gillin, Y. Li, J. Y. Chang, et al. Motion-robust intensity-modulated proton therapy for distal esophageal cancer. *Med. Phys.*, 43(3):1111–1118, 2016.

[72] S. van de Water, I. van Dam, D. R. Schaart, A. Al-Mamgani, B. J. Heijmen, and M. S. Hoogeman. The price of robustness; impact of worst-case optimization on organ-at-risk dose and complication probability in intensity-modulated proton therapy for oropharyngeal cancer patients. *Radiother. Oncol.*, 120(1):56–62, 2016.

[73] A. Van de Schoot, J. Visser, Z. Van Kesteren, T. Janssen, C. Rasch, and A. Bel. Beam configuration selection for robust intensity-modulated proton therapy in cervical cancer using pareto front comparison. *Phys. Med. Biol.*, 61(4):1780, 2016.

[74] J. Shan, Y. An, M. Bues, S. E. Schild, and W. Liu. Robust optimization in IMPT using quadratic objective functions to account for the minimum MU constraint. *Med. Phys.*, 2018.

[75] J. Steitz, P. Naumann, S. Ulrich, M. F. Haefner, F. Sterzing, U. Oelfke, and M. Bangert. Worst case optimization for interfractional motion mitigation in carbon ion therapy of pancreatic cancer. *Radiation Oncology*, 11(1):134, 2016.

[76] M. Lowe, A. Aitkenhead, F. Albertini, A. J. Lomax, and R. I. MacKay. A robust optimisation approach accounting for the effect of fractionation on setup uncertainties. *Phys. Med. Biol.*, 62(20):8178, 2017.

[77] K. Bernatowicz, X. Geets, A. Barragan, G. Janssens, K. Souris, and E. Sterpin. Feasibility of online IMPT adaptation using fast, automatic and robust dose restoration. *Phys. Med. Biol.*, 63(8):085018, 2018.

[78] A. Tasson, N. N. Laack, and C. Beltran. Clinical implementation of robust optimization for craniospinal irradiation. *Cancers*, 10(1):7, 2018.

[79] T. Arts, S. Breedveld, M. A. de Jong, E. Astreinidou, L. Tans, F. Keskin-Cambay, A. D. Krol, S. van de Water, R. G. Bijman, and M. S. Hoogeman. The impact of treatment accuracy on proton therapy patient selection for oropharyngeal cancer patients. *Radiother. Oncol.*, 125(3):520–525, 2017.

[80] C. Liu, S. E. Schild, J. Y. Chang, Z. Liao, S. Korte, J. Shen, X. Ding, Y. Hu, Y. Kang, S. R. Keole, et al. Impact of spot size and spacing on the quality of robustly optimized intensity modulated proton therapy plans for lung cancer. *Int. J. Radiat. Oncol. Biol. Phys.*, 101(2):479–489, 2018.

[81] M. A. van de Sande, C. L. Creutzberg, S. van de Water, A. W. Sharfo, and M. S. Hoogeman. Which cervical and endometrial cancer patients will benefit most from intensity-modulated proton therapy? *Radiother. Oncol.*, 120(3):397–403, 2016.

[82] R. G. Bijman, S. Breedveld, T. Arts, E. Astreinidou, M. A. de Jong, P. V. Granton, S. F. Petit, and M. S. Hoogeman. Impact of model and dose uncertainty on model-based selection of oropharyngeal cancer patients for proton therapy. *Acta Oncologica*, 56(11):1444–1450, 2017.

[83] J. Y. Chang, H. Li, X. R. Zhu, Z. Liao, L. Zhao, A. Liu, Y. Li, N. Sahoo, F. Poenisch, D. R. Gomez, R. Wu, M. Gillin, and X. Zhang. Clinical implementation of intensity modulated

proton therapy for thoracic malignancies. *Int. J. Radiat. Oncol. Biol. Phys.*, 90(4):809–818, 2014.

[84] Z. Morávek, M. Rickhey, M. Hartmann, and L. Bogner. Uncertainty reduction in intensity modulated proton therapy by inverse Monte Carlo treatment planning. *Phys. Med. Biol.*, 54(15):4803–4819, 2009.

[85] T. Inaniwa, N. Kanematsu, T. Furukawa, and A. Hasegawa. A robust algorithm of intensity modulated proton therapy for critical tissue sparing and target coverage. *Phys. Med. Biol.*, 56(15):4749–4770, 2011.

[86] T. Inaniwa, N. Kanematsu, T. Furukawa, and K. Noda. Optimization algorithm for overlapping-field plans of scanned ion beam therapy with reduced sensitivity to range and setup uncertainties. *Phys. Med. Biol.*, 56(6):1653–1669, 2011.

[87] W. Cao, G. J. Lim, A. Lee, Y. Li, W. Liu, X. R. Zhu, and X. Zhang. Uncertainty incorporated beam angle optimization for IMPT treatment planning. *Med. Phys.*, 39(8):5248–5256, 2012.

[88] A. Fredriksson. A characterization of robust radiation therapy treatment planning methods – from expected value to worst case optimization. *Med. Phys.*, 39(8):5169–5181, 2012.

[89] W. Liu, Y. Li, X. Li, W. Cao, and X. Zhang. Influence of robust optimization in intensity-modulated proton therapy with different dose delivery techniques. *Med. Phys.*, 39(6):3089–3101, 2012.

[90] M. Bangert, P. Hennig, and U. Oelfke. Analytical probabilistic modeling for radiation therapy treatment planning. *Phys. Med. Biol.*, 58(16):5401, 2013.

[91] W. Liu, S. J. Frank, X. Li, Y. Li, P. C. Park, L. Dong, X. Ronald Zhu, and R. Mohan. Effectiveness of robust optimization in intensity-modulated proton therapy planning for head and neck cancers. *Med. Phys.*, 40(5), 2013.

[92] S. Petit, J. Seco, and H. Kooy. Increasing maximum tumor dose to manage range uncertainties in IMPT treatment planning. *Phys. Med. Biol.*, 58(20):7329–7341, 2013.

[93] A. Fredriksson, A. Forsgren, and B. Hårdemark. Maximizing the probability of satisfying the clinical goals in radiation therapy treatment planning under setup uncertainty. *Med. Phys.*, 42(7):3992–3999, 2015.

[94] W. Liu, R. Mohan, P. Park, Z. Liu, H. Li, X. Li, Y. Li, R. Wu, N. Sahoo, L. Dong, et al. Dosimetric benefits of robust treatment planning for intensity modulated proton therapy for base-of-skull cancers. *Practical radiation oncology*, 4(6):384–391, 2014.

[95] A. Fredriksson and R. Bokrantz. A critical evaluation of worst case optimization methods for robust intensity-modulated proton therapy planning. *Med. Phys.*, 41(8), 2014.

[96] W. Liu, Z. Liao, S. E. Schild, Z. Liu, H. Li, Y. Li, P. C. Park, X. Li, J. Stoker, J. Shen, et al. Impact of respiratory motion on worst-case scenario optimized intensity modulated proton therapy for lung cancers. *Practical radiation oncology*, 5(2):e77–e86, 2015.

[97] W. Liu, S. E. Schild, J. Y. Chang, Z. Liao, Y.-H. Chang, Z. Wen, J. Shen, J. B. Stoker, X. Ding, Y. Hu, et al. Exploratory study of 4D versus 3D robust optimization in intensity modulated proton therapy for lung cancer. *Int. J. Radiat. Oncol. Biol. Phys.*, 95(1):523–533, 2016.

[98] R. Bokrantz and A. Fredriksson. Scenario-based radiation therapy margins for patient setup, organ motion, and particle range uncertainty. *Phys. Med. Biol.*, 62(4):1342, 2017.

[99] N. Wahl, P. Hennig, H. Wieser, and M. Bangert. Efficiency of analytical and sampling-based uncertainty propagation in intensity-modulated proton therapy. *Phys. Med. Biol.*, 62(14):5790, 2017.

[100] N. Wahl, P. Hennig, H.-P. Wieser, and M. Bangert. Analytical incorporation of fractionation effects in probabilistic treatment planning for intensity-modulated proton therapy. *Med. Phys.*, 2018.

[101] A. Fredriksson and R. Bokrantz. The scenario-based generalization of radiation therapy margins. *Phys. Med. Biol.*, 61(5):2067, 2016.

[102] P. C. Park, J. Cheung, X. R. Zhu, N. Sahoo, L. Dong, et al. Fast range-corrected proton dose approximation method using prior dose distribution. *Phys. Med. Biol.*, 57(11):3555, 2012.

[103] H. Wieser, P. Hennig, N. Wahl, and M. Bangert. Analytical probabilistic modeling of RBE-weighted dose for ion therapy. *Phys. Med. Biol.*, 62(23):8959, 2017.

[104] M. Söhn, M. Weinmann, and M. Alber. Intensity-modulated radiotherapy optimization in a quasi-periodically deforming patient model. *Int. J. Radiat. Oncol. Biol. Phys.*, 75(3):906–914, 2009.

[105] P. Zhang, G. D. Hugo, and D. Yan. Planning study comparison of real-time target tracking and four-dimensional inverse planning for managing patient respiratory motion. *Int. J. Radiat. Oncol. Biol. Phys.*, 72(4):1221–1227, 2008.

[106] W. T. Watkins, J. A. Moore, J. Gordon, G. D. Hugo, and J. V. Siebers. Multiple anatomy optimization of accumulated dose. *Medical Physics*, 41(11):111705, 2014.

[107] E. Lens, A. N. Kotte, A. Patel, H. D. Heerkens, M. Bal, G. van Tienhoven, A. Bel, A. van der Horst, and G. J. Meijer. Probabilistic treatment planning for pancreatic cancer treatment: prospective incorporation of respiratory motion shows only limited dosimetric benefit. *Acta Oncologica*, 56(3):398–404, 2017.

[108] T. Bortfeld, T. C. Chan, A. Trofimov, and J. N. Tsitsiklis. Robust management of motion uncertainty in intensity-modulated radiation therapy. *Operations Research*, 56(6):1461–1473, 2008.

[109] J. Unkelbach. *Inclusion of organ motion in IMRT optimization using probabilistic treatment planning*. PhD thesis, University of Heidelberg, 2006. Available at https://katalog.ub.uni-heidelberg.de/titel/66118919.

[110] T. C. Chan, H. Mahmoudzadeh, and T. G. Purdie. A robust-CVaR optimization approach with application to breast cancer therapy. *European Journal of Operational Research*, 238(3):876–885, 2014.

[111] H. Mahmoudzadeh, J. Lee, T. C. Chan, and T. G. Purdie. Robust optimization methods for cardiac sparing in tangential breast IMRT. *Med. Phys.*, 42(5):2212–2222, 2015.

[112] H. Mahmoudzadeh, T. G. Purdie, and T. C. Chan. Constraint generation methods for robust optimization in radiation therapy. *Operations Research for Health Care*, 8:85–90, 2016.

[113] T. C. Chan, J. N. Tsitsiklis, and T. Bortfeld. Optimal margin and edge-enhanced intensity maps in the presence of motion and uncertainty. *Phys. Med. Biol.*, 55(2):515, 2010.

[114] T. C. Chan. Motion-compensating intensity maps in intensity-modulated radiation therapy. *IIE Transactions on Healthcare Systems Engineering*, 3(1):1–22, 2013.

[115] C. Vrančić, A. Trofimov, T. C. Chan, G. C. Sharp, and T. Bortfeld. Experimental evaluation of a robust optimization method for IMRT of moving targets. *Phys. Med. Biol.*, 54(9):2901, 2009.

[116] C. McCann, T. Purdie, A. Hope, A. Bezjak, and J.-P. Bissonnette. Lung sparing and dose escalation in a robust-inspired IMRT planning method for lung radiotherapy that accounts for intrafraction motion. *Med. Phys.*, 40(6), 2013.

[117] M. Ahanj, J.-P. Bissonnette, E. Heath, and C. McCann. Robustness assessment of a novel IMRT planning method for lung radiotherapy. *Physica Medica*, 32(6):749–757, 2016.

[118] D. Yan, F. Vicini, J. Wong, and A. Martinez. Adaptive radiation therapy. *Phys. Med. Biol.*, 42(1):123, 1997.

[119] T. C. Chan and V. V. Mišić. Adaptive and robust radiation therapy optimization for lung cancer. *European Journal of Operational Research*, 231(3):745–756, 2013.

[120] V. V. Mišić and T. C. Chan. The perils of adapting to dose errors in radiation therapy. *PloS one*, 10(5):e0125335, 2015.

[121] P. A. Mar and T. C. Chan. Adaptive and robust radiation therapy in the presence of drift. *Phys. Med. Biol.*, 60(9):3599, 2015.

[122] A. Trofimov, E. Rietzel, H. Lu, B. Martin, S. Jiang, G. T. Y. Chen, and T. Bortfled. Temporo-spatial IMRT optimization: concepts, implementation and initial results. *Phys. Med. Biol.*, 50:2779–98, 2005.

[123] O. Nohadani, J. Seco, and T. Bortfeld. Motion management with phase-adapted 4D-optimization.

Phys. Med. Biol., 55(17):5189, 2010.

[124] Y. Suh, W. Murray, and P. J. Keall. IMRT treatment planning on 4D geometries for the era of dynamic MLC tracking. *Technology in cancer research & treatment*, 13(6):505–515, 2014.

[125] E. Chin and K. Otto. Investigation of a novel algorithm for true 4D-VMAT planning with comparison to tracked, gated and static delivery. *Med. Phys.*, 38(5):2698–2707, 2011.

[126] Y. Ma, D. Chang, P. Keall, Y. Xie, J.-y. Park, T.-S. Suh, and L. Xing. Inverse planning for four-dimensional (4D) volumetric modulated arc therapy. *Med. Phys.*, 37(11):5627–5633, 2010.

[127] K. Bernatowicz, Y. Zhang, R. Perrin, D. C. Weber, and A. J. Lomax. Advanced treatment planning using direct 4D optimisation for pencil-beam scanned particle therapy. *Phys. Med. Biol.*, 62(16):6595, 2017.

[128] C. Bert, C. Graeff, M. Riboldi, S. Nill, G. Baroni, and A.-C. Knopf. Advances in 4D treatment planning for scanned particle beam therapyreport of dedicated workshops. *Technology in cancer research & treatment*, 13(6):485–495, 2014.

[129] E. Engwall, A. Fredriksson, and L. Glimelius. 4D robust optimization including uncertainties in time structures can reduce the interplay effect in proton pencil beam scanning radiation therapy. *Med. Phys.*, 45(9):4020–4029, 2018.

[130] K. Jordan. Robust planning techniques using Pinnacle 3 TPS IMPT module. *Med. Phys.*, 44(6):3080, 2017.

[131] L. C. Goddard, N. P. Brodin, W. R. Bodner, M. K. Garg, and W. A. Tomé. Comparing photon and proton-based hypofractioned sbrt for prostate cancer accounting for robustness and realistic treatment deliverability. *British Journal of Radiology*, 91(1085):20180010, 2018.

[132] K. R. Jethwa, E. J. Tryggestad, T. J. Whitaker, B. T. Giffey, B. D. Kazemba, M. A. Neben-Wittich, K. W. Merrell, M. G. Haddock, and C. L. Hallemeier. Initial experience with intensity modulated proton therapy for intact, clinically-localized pancreas cancer: clinical implementation, dosimetric analysis, acute treatment-related adverse events, and patient-reported outcomes. *Advances in Radiation Oncology*, 2018.

[133] H. Miura, S. Ozawa, and Y. Nagata. Efficacy of robust optimization plan with partial-arc VMAT for photon volumetric-modulated arc therapy: A phantom study. *Journal of Applied Clinical Medical Physics*, 18(5):97–103, 2017.

[134] M. Byrne, Y. Hu, and B. Archibald-Heeren. Evaluation of RayStation robust optimisation for superficial target coverage with setup variation in breast IMRT. *Australasian physical & engineering sciences in medicine*, 39(3):705–716, 2016.

[135] C. A. Jensen, A. M. A. Roa, M. Johansen, J.-A. Lund, and J. Frengen. Robustness of VMAT and 3DCRT plans toward setup errors in radiation therapy of locally advanced left-sided breast cancer with dibh. *Physica Medica*, 45:12 – 18, 2018.

[136] X. Zhang, Y. Rong, S. Morrill, J. Fang, G. Narayanasamy, E. Galhardo, S. Maraboyina, C. Croft, F. Xia, and J. Penagaricano. Robust optimization in lung treatment plans accounting for geometric uncertainty. *Journal of applied clinical medical physics*, 2018.

[137] B. R. Archibald-Heeren, M. V. Byrne, Y. Hu, M. Cai, and Y. Wang. Robust optimization of VMAT for lung cancer: Dosimetric implications of motion compensation techniques. *Journal of applied clinical medical physics*, 18(5):104–116, 2017.

[138] A.-K. Exeli, D. Kellner, L. Exeli, P. Steininger, F. Wolf, F. Sedlmayer, and H. Deutschmann. Cerebral cortex dose sparing for glioblastoma patients: IMRT versus robust treatment planning. *Radiation Oncology*, 13(1):20, 2018.

[139] A. Ramlow, M. S. Assenholt, M. F. J. amd C. Grønborg, R. Nout, M. Alber, L. Fokdal, K. Tanderup, and J. C. Lindegaard. Clinical implementation of coverage probability planning for nodal boosting in locally advanced cervical cancer. *Radiother. Oncol.*, 123(1):158–163, 2017.

[140] J. C. Lindegaard, M. S. Assenholt, A. Ramlow, L. Fokdal, M. Alber, and K. Tanderup. Early clinical outcome of coverage probability based treatment planning in locally advanced cervical cancer for simultaneous integrated boost of nodes. *Acta Oncologica*, 56(11):1479–1486, 2017.

[141] R. Pötter, K. Tanderup, C. Kirisits, A. de Leeuw, K. Kirchheiner, R. Nout, L. T. Tan, C. Haie-Meder, U. Mahantshetty, B. Segedin, et al. The EMBRACE II study: The outcome and prospect of two decades of evolution within the GEC-ESTRO GYN working group and the EMBRACE studies. *Clin Transl Radiat Oncol*, pages 48–60, 2018.